O estatuto científico da ciência cognitiva em sua fase inicial

uma análise a partir da estrutura das revoluções científicas de Thomas Kuhn

Marcos Antonio Alves
Alan Rafael Valente

O estatuto científico da ciência cognitiva em sua fase inicial:

uma análise a partir da estrutura das revoluções científicas de Thomas Kuhn

Marcos Antonio Alves
Alan Rafael Valente

O estatuto científico da ciência cognitiva em sua fase inicial:

uma análise a partir da estrutura das revoluções científicas de Thomas Kuhn

Marília/Oficina Universitária
São Paulo/Cultura Acadêmica

2021

Sumário

Apresentação

Esta obra foi publicada a partir de edital interno de publicação de trabalhos de docentes e egressos do Programa de Pós-Graduação em Filosofia (PPGFIL) da Unesp. Situado no *campus* de Marília, o PPGFIL iniciou suas atividades em 1996. Trata-se de um programa consolidado que apresenta bons resultados em diferentes âmbitos. São dignas de nota a quantidade e a qualidade das publicações de seus docentes, discentes e egressos. Atividades de ensino, pesquisa e extensão, inserção social, internacionalização, bem como a formação de novos quadros para a filosofia também são marcantes. Já são tradicionais e de grande visibilidade, por exemplo, alguns eventos promovidos e realizados pelo programa. Já são mais de 250 egressos, muitos deles concursados nas redes estaduais de ensino básico ou em instituições de ensino superior em todo o país. Boa parte deles cursou doutorado, realizou estágio ou pesquisa em instituições nacionais e estrangeiras de renome.

Como parte das comemorações de seu jubileu de prata, o PPGFIL vem realizando e promovendo uma série de atividades em diversos segmentos. Em uma frente, vem reestruturando suas linhas de pesquisa, seu corpo docente, bem como seus projetos e grupos de pesquisa. Em relação às linhas, em 2020 elas passaram a ser apenas duas, intituladas "Filosofia da Informação, da Cognição e da Consciência" e "Conhecimento, Ética

https://doi.org/10.36311/2021.978-65-5954-052-5.p7-12

e Política". Tais modificações buscam manter e respeitar a liberdade, a autonomia e a visão filosófica dos grupos ou dos integrantes do programa.

Com as mudanças, resultado de seu processo de autoavaliação, o programa reuniu docentes em torno de temas e pesquisas convergentes. Com isso, visa a favorecer o desenvolvimento ainda mais substancial e aprofundado de pesquisas, produzindo conhecimento qualificado, ampliando a internacionalização, melhorando a formação de seus discentes, a inserção social através da socialização do conhecimento, realização de eventos, desenvolvimento de projetos de ensino, pesquisa e extensão.

O programa está solicitando, depois de um longo trabalho coletivo, a abertura do seu doutorado. O curso pretende atender à demanda de discentes formados na graduação e mestrado em filosofia e em outros cursos da própria Unesp, além de estudantes oriundos de diversas regiões do país interessados em aprofundar suas pesquisas nos temas e problemas abordados no PPGFIL. Com isso, favorecerá a formação continuada de discentes na Unesp, da graduação ao doutorado, acolhendo também candidatos de outras instituições interessados em desenvolver pesquisas nas áreas de especialidade de seus docentes.

Em outra frente, o programa reformulou e intensificou sua interação com a comunidade por meio das redes sociais. Por meio das publicações em sua página no Facebook, no endereço https://www.facebook.com/posfilmarilia, deixa os seguidores informados das suas atividades. Já a sua página oficial está hospedada no site da FFC/Unesp/*Campus* de Marília, que pode ser acessada no endereço http://www.marilia.unesp.br/posfil. Além de publicações sobre sua atividade cotidiana, oferece variadas informações referentes a seu histórico, missão, objetivos, processo seletivo, bem como possui seções especificamente direcionadas a discentes, docentes e egressos. Buscando melhor comunicação, acessibilidade e transparência, a página, depois de reformulada, está mais leve, informativa e acessível.

A socialização do conhecimento e contato com a comunidade também é efetivada através das revistas científicas vinculadas ao programa. Dentre elas, estão a *Kínesis*: Revista dos Estudos dos Pós-Graduandos em Filosofia, e a *Trans/Form/Ação*: revista de filosofia da Unesp, já considerada patrimônio do curso de filosofia da Unesp e um dos mais conceituados periódicos na área tanto no Brasil quanto no exterior. A *Kínesis*, como

diz o próprio sobrenome, é voltada principalmente, mas não somente, à publicação de trabalhos de pós-graduandos. Já a *Trans/Form/Ação* publica textos de profissionais em filosofia e áreas afins. Ambas são voltadas à publicação de trabalhos de filosofia ou de interesse filosófico, difundindo o conhecimento produzido na área tanto no Brasil quanto no exterior.

Ainda como parte da comemoração dos seus 25 anos, o PPGFIL lançou o edital para publicação de livros de docentes e egressos, ao qual este livro foi submetido e aprovado para publicação. As propostas submetidas foram avaliadas na plataforma da revista *Trans/Form/Ação*, no caráter de parecer duplo-cego. Tal acordo de cooperação foi pensado para garantir transparência e confiabilidade no processo seletivo das submissões. Ao receber a solicitação de avaliação, os pareceristas também foram convidados a produzir o prefácio do livro, caso deliberassem pela aprovação da obra. Além de favorecer ainda mais o cuidado no trabalho avaliativo, com essa atitude buscamos valorizar ainda mais a contribuição dos avaliadores.

As obras aprovadas no edital foram publicadas em conjunto pelas editoras Oficina Universitária e Cultura Acadêmica. A Oficina Universitária é um selo editorial da Faculdade de Filosofia e Ciências da Unesp, *campus* de Marília, apoiada pelo Laboratório Editorial da FFC. Foi instituída com o objetivo de criar condições e oportunidades para a difusão de pesquisas e tornar públicos os resultados dos trabalhos do corpo docente da FFC. Já a Cultura Acadêmica, selo da Fundação Editora da Unesp, visa auxiliar principalmente o atendimento às múltiplas demandas editoriais da Unesp. Com a ampliação do número de títulos editados pelo selo, são abertas novas oportunidades de publicação num momento em que a pesquisa acadêmica e sua divulgação são cada vez mais necessárias.

É com grande prazer e satisfação que publicamos este livro, intitulado *O estatuto científico da ciência cognitiva em sua fase inicial: uma análise a partir da estrutura das revoluções científicas de Thomas Kuhn*, escrito por Marcos Antonio Alves e Alan Rafael Valente e prefaciado por João de Fernandes Teixeira.

Os autores analisam o estatuto científico da ciência cognitiva nos seus primeiros anos, mais especificamente entre as décadas de 1940 e 1970, a partir do aparato conceitual da estrutura das revoluções científicas de Thomas Kuhn. Eles dividem a obra em três capítulos. No primeiro, como ponto de partida para a análise, apresentam a abordagem kuhniana

do progresso da ciência. Normalmente, a fase inicial de uma nova área de pesquisa é marcada por um momento de luta paradigmática, caracterizando-se pela rivalidade entre paradigmas. Uma vez superado este momento de pré-ciência, a área entra em um período de ciência normal, marcado pela prevalência de um paradigma. Em um sentido geral, lembram os autores, Kuhn trata de mudanças significativas de paradigmas geradas por momentos de crise que alteram radicalmente uma disciplina e, de certa forma, a ciência em um sentido geral. Em outra direção, mas complementar a esse processo revolucionário, Kuhn trata do progresso interno de uma área de pesquisa, caracterizada pelo aprofundamento de teorias, metodologias, conceitos, experimentos e observações empíricas, fortalecimento de uma comunidade científica em volta do paradigma dominante. O objetivo dos autores neste livro é averiguar o estatuto da ciência cognitiva em sua formação. Teria ela começado como pré-ciência ou já teria se estabelecido como ciência normal? Para responder a essa questão, apresentam o surgimento dessa área de pesquisa em dois contextos: histórico e epistemológico. O aspecto histórico do seu surgimento é exposto no segundo capítulo, focalizando o ambiente científico do momento, os primeiros textos que fazem referência a ela, a formação dos primeiros membros, encontros e temas que serviram como base para a sua emergência. No terceiro capítulo eles expõem o aspecto epistemológico, ou seja, o arcabouço conceitual, metodológico, metafísico e teórico das pesquisas na área, bem como as suas duas principais vertentes: o cognitivismo e o conexionismo. Feito isso, com base nestes elementos, Alves e Valente finalizam o livro expressando o estatuto da ciência cognitiva no momento averiguado, com base na abordagem de Kuhn.

Este e os demais livros publicados por este edital podem ser baixados gratuitamente no catálogo da editora Oficina Universitária: https://ebooks.marilia.unesp.br/index.php/lab_editorial. São eles:

- Eichmann e a incapacidade de pensar: alienação do mundo e do pensamento em Hannah Arendt. Renato de Oliveira Pereira
- Hábitos motores e identidade pessoal. Ana Paula Talin Bissoli
- O estatuto científico da ciência cognitiva em sua fase inicial: uma análise a partir da estrutura das revoluções científicas de Thomas Kuhn. Marcos Antonio Alves e Alan Rafael Valente

- Semiótica e Pragmatismo. Interfaces teóricas. Vol. I. Ivo Assad Ibri
- Semiótica e Pragmatismo. Interfaces teóricas. Vol. II. Ivo Assad Ibri
- Verdade e arte: a concepção ontológica da obra de arte no pensamento de Martin Heidegger. Juliano Rabello

Esperamos, com esta atividade, fazer cumprir um dos objetivos de um programa de pós-graduação, o de produzir e socializar o conhecimento. Desejamos aos leitores desta e das demais obras uma reflexão profícua oriunda de sua leitura.

Marcos Antonio Alves

Coordenador do Programa de Pós-Graduação em Filosofia da Unesp

Prefácio

Prefaciar um livro tão instigante é um desafio. Nele, os autores buscam discutir uma questão fundamental: o estatuto científico da ciência cognitiva. Será que a ciência cognitiva merece, efetivamente, ser considerada uma ciência?

Para responder a essa questão, eles recorrem a ferramentas teóricas e conceituais da epistemologia histórica de Thomas Kuhn. A proposta é analisar a produção da ciência cognitiva entre as décadas de 1940 e 1970 para saber se nesse período se consolidou um paradigma unificador para essa nova disciplina que ensaia seus primeiros passos – um paradigma que abranja os diversos tipos de investigação nesse período e liberte a ciência cognitiva de um estado pré-paradigmático, no qual não existe concordância sequer com os pressupostos e princípios básicos que essa nova disciplina deveria seguir.

Terá a ciência cognitiva, nesse período de 1940 a 1970, ultrapassado a fase pré-paradigmática e alcançado um paradigma a partir do qual seja possível identificar um período de ciência normal? A resposta dos autores é negativa, apesar de ser possível identificar um momento de concordância, um assentimento transitório sobre alguns princípios norteadores de uma ciência da mente compartilhados pelos cientistas cognitivos. A esse momento os autores chamam de "calmaria", no que

https://doi.org/10.36311/2021.978-65-5954-052-5.p13-16

antes parecia ser uma guerra de todos contra todos. Mas a calmaria é um momento efêmero. Há uma dispersão radical que atinge profundamente os movimentos internos da ciência cognitiva; uma discordância acerca do que deveria ser uma ciência da mente. Daí a incapacidade de consolidar um paradigma.

A análise histórica apresentada pelos autores é rigorosa, precisa. A ciência cognitiva não passa pelo crivo da epistemologia histórica. No entanto, terão outros movimentos, na longa história das tentativas de consolidar uma ciência da mente, conseguido uma consolidação paradigmática? Não podemos confundir certa consolidação institucional com a maturidade paradigmática. É o que ocorre com a psicanálise, o behaviorismo e outros movimentos na história das tentativas de formular uma ciência da mente, os quais, até hoje, tentam se impor institucionalmente, mas não conseguem uma univocidade entre seus adeptos que lhes permita superar sua dispersão interna. Não há um paradigma unificador da psicanálise, pois ela não é um movimento monolítico. Há discordâncias quanto a pressupostos básicos. Isso também ocorre com o behaviorismo. Na psicanálise há os junguianos, os bionianos, os freudianos e os lacanianos. Há behavioristas radicais (skinnerianos), behavioristas watsonianos, behavioristas cognitivos e assim por diante. Haverá algum paradigma unificador para a psicanálise e para o behaviorismo? Como conciliar Freud e Jung, ou Watson e Skinner? Visto dessa maneira, talvez o problema não seja com a ciência cognitiva, mas com as tentativas de fundar uma ciência da mente que, até agora, não foram bem-sucedidas.

Talvez a filosofia possa fornecer um guarda-chuva para abrigar a dispersão da ciência cognitiva. Penso que uma raiz comum para todos os movimentos da ciência cognitiva é a proposta de epistemologia naturalizada formulada pelo filósofo Willard van Orman Quine. Ele defendeu a naturalização da epistemologia, ou seja, a ideia de que o estudo do conhecimento não cabe à filosofia, mas à ciência. A ciência deve se debruçar sobre si mesma para compreender como o conhecimento é produzido. A ciência que pode debruçar sobre si mesma e compreender como funciona nosso conhecimento é a ciência cognitiva. Ou melhor dito, a ciência da cognição. Talvez tenhamos de nos resignar a esse único guarda-chuva proposto por Quine e a abandonar a pretensão de elaborar uma ciência cognitiva que possa desfrutar de um paradigma, uma definição

de um ponto de partida que, assumido por vários pesquisadores, leve ao desenvolvimento do que Thomas Kuhn chamou de ciência normal, ou seja, períodos históricos em que não há questionamento de pressupostos por parte dos cientistas.

Como afirmei, o livro aborda um trecho específico da história da ciência cognitiva, de 1940 a 1970. Chamo esse intervalo de época pós-Turing. Um período de grande importância para a ciência cognitiva, apesar dela ter seguido um caminho errático. Atualmente, a dispersão continua. Há, de um lado, a neurociência cognitiva, que teve grande avanço após a invenção da neuroimagem na metade da década de 1990. A neurociência cognitiva pretende monopolizar o que entendemos por ciência cognitiva e reduzi-la à neurociência. Por outro lado, há a inteligência artificial, baseada nas redes neurais, nos algoritmos de aprendizado e no *big data*. Os defensores das redes neurais modelam o cérebro como uma máquina estatística. Essas abordagens coexistem e disputam a primazia de ser a única e definitiva ciência da mente. Continuamos em um estado pré-paradigmático. Parece que atualmente a única concordância é o abandono do modelo computacional da mente que foi definitivamente sepultado. Ninguém acredita mais que o cérebro seja um computador ou que a mente seja o *software* do cérebro.

Recomendo vivamente a leitura deste livro. Ele é uma abordagem inédita da história da ciência cognitiva, que, em poucas décadas de existência, continua refletindo uma profunda dissensão entre os modos de abordar uma ciência da mente. Mais do que uma reflexão epistemológica, ele traz, também, uma contribuição para a história da ciência cognitiva no século passado, em linguagem fluente e agradável.

João de Fernandes Teixeira

Pesquisador, filósofo e docente aponsentado na Universidade Federal de São Carlos

Introdução

O período entre a segunda metade do século XIX e o começo do XX testemunhou a proposta de diferentes abordagens explicativas de fenômenos do mundo consideradas revolucionárias. Dentre elas, citamos o evolucionismo de Darwin (1809-1882), o socialismo de Marx (1818-1883), a psicanálise de Freud (1856-1939), a relatividade de Einstein (1879-1955). Essas novas perspectivas produziram mudanças em diversos âmbitos, como alteração de visão de mundo, descoberta de novos fenômenos e elementos naturais, abordagens metodológicas e técnicas inovadoras para tratar dos problemas investigados na ciência.

Nos anos 1940 começou a despontar outro projeto investigativo que também acabou sendo revolucionário. Embalados no movimento que acabou sendo denominado Virada Informacional, emergiram muitas pesquisas buscando tratar de um tipo especial de fenômeno do mundo: os processos cognitivos.

Embora recente, esse objeto de pesquisa não é novo. Desde os primórdios da filosofia, muitos pensadores, cada um à sua forma, já se debruçaram sobre questões semelhantes às desse auspicioso movimento intelectual. Salvaguardando os seus limites e as suas peculiaridades históricas, o tema da cognição sempre foi tomado como objeto de estudo dos filósofos, como podemos observar, por exemplo, em *Fédon*, de Platão,

De anima, de Aristóteles, e *Meditações Metafísicas*, de Descartes, apenas para ilustrar alguns deles.

Os pesquisadores contemporâneos, no entanto, não pretendiam adotar um método discursivo ou calcado na metafísica, como de costume na filosofia. Visavam, de um modo ou de outro, poder observar, manipular, criar teorias empiricamente testáveis dos processos cognitivos. Muitos desejavam adotar um método empírico, algumas vezes semelhante ao da Física em suas pesquisas. Dentre as diversas áreas envolvidas neste trabalho, podemos citar biologia, psicologia, física, linguística, neurociências, filosofia, cibernética e computação. Tais ciências foram denominadas ciências cognitivas.

Tais ciências se caracterizam pelo fato de pesquisar a cognição cada uma a seu modo. No início das suas atividades, existia a preocupação por parte de alguns pesquisadores em fazer com que a interação entre cientistas de áreas diferentes criasse um campo de pesquisa unificado e um vocabulário comum para a explicação dos processos cognitivos humanos. Na visão idealizada desses pesquisadores, todos os cientistas das ciências cognitivas deveriam adequar-se a um mesmo método e vocabulário. Ainda assim, nem por isso as ciências cognitivas perderiam sua identidade. Seus integrantes continuariam mantendo um diálogo com estudiosos de outras áreas sem deixar de ter seu modo particular de pesquisa.

Uma das ciências cognitivas é a ciência cognitiva. Sua característica principal é a ênfase dada à função exercida pelos computadores no estudo da mente, qual seja, o de fornecer estudos explicativos e preditivos de processos cognitivos. Ela nasce com o intuito de responder empiricamente, através do uso do computador, questões que os primeiros filósofos já haviam colocado a respeito de aspectos cognitivos. Além de uma tendência interdisciplinar, essa ciência possui, ou pretende possuir, método e vocabulário próprios no que tange à sua atividade.

Diante de tantos novos programas de pesquisa, com pretensas teorias, um dos propósitos centrais de epistemólogos, especialmente no começo do século XX, consistia em saber quais deles realmente produziam conhecimento, quais eram capazes de oferecer explicações e previsões aos fenômenos no mundo, quais poderiam ser considerados científicos. A busca por um critério de demarcação científica, distinguindo a ciência da

não ciência, tornou-se uma das questões centrais da filosofia da ciência no século passado, mantendo-se ainda em certa evidência atualmente.

Neste livro desenvolvemos uma análise do estatuto científico da ciência cognitiva nos seus primeiros anos, mais especificamente entre as décadas de 1940 e 1970. Seria tal área de pesquisa considerada científica? Em que medida suas teorias dizem algo a respeito do mundo? Ela teria começado como uma área conflituosa ou já iniciou com um grupo unido e coeso em direção uníssona para a resolução de problemas e propostas de teorias explicativas de certos fenômenos do mundo? Houve algum progresso na fase inicial de suas atividades?

A fim de tratar da demarcação e de outras questões, como a do progresso e método científicos, por exemplo, surgiram perspectivas epistemológicas como o Empirismo Lógico, o Falsificacionismo e o Estruturalismo. Esta última, particularmente, possui diferentes versões, dentre elas a de Imre Lakatos (1922-1974) e a de Thomas Kuhn (1922-1996).

No começo do século XX, pesquisadores de diferentes áreas se reuniam em Viena, local de efervescência cultural, científica e filosófica, com o objetivo de discutir questões como as expostas acima, fundando uma perspectiva denominada Empirismo Lógico ou Positivismo Lógico. Dentre esses pensadores figuravam o físico Moritz Schlick, os matemáticos Hans Hahn e Rudolf Carnap e o sociólogo e economista Otto Neurath. De modo geral, para os integrantes do Círculo de Viena, a ciência começa com a observação. A observação é neutra, imparcial, ou seja, genuína, e fornece uma base segura para a ciência. Uma observação é explicitada ou expressa através de um enunciado elementar ou protocolar, também chamado de proposição de observação. Ele descreve experiências empíricas elementares, ou seja, uma observação de um indivíduo, num espaço e tempo específicos. Um número adequado de observações sobre um fenômeno sob uma ampla variedade de condições, todas satisfazendo o referido fenômeno, justificaria a sua generalização para casos gerais, aplicando-o a todos os casos do mesmo tipo. Assim, por exemplo, após observar muitos cisnes, sob uma ampla variedade de condições, tais como em lugares e momentos diferentes, e averiguar que cada um deles possui a cor branca, admite-se generalizar legitimamente que "todo cisne é branco". Por esse processo indutivo, ocorre a descoberta de leis científicas. Nessa perspectiva, a ciência

se desenvolve a partir do acúmulo de dados: quanto maior o número de observações, mais força possui uma lei científica.

Muitas críticas foram dirigidas a essa perspectiva. Uma delas refere-se à possibilidade de a observação ser genuína. A observação, conforme aponta Hanson (1975), geralmente é guiada por interpretações, valores, expectativas, preconceitos e crenças do observador. A indução, inferência basilar do indutivismo lógico, também sofreu diversas críticas, dada a impossibilidade lógica e empírica da sua validade. Observar, por exemplo, um grande número de cisnes, por maior que seja a amostragem, não garante que todos os cisnes sejam brancos. É possível o surgimento de um cisne não branco que refute o enunciado geral, ou seja, falsifique a conclusão da inferência, ainda que as premissas se mantenham verdadeiras.

Um dos críticos do Círculo de Viena é o filósofo austríaco Karl Popper. Ele desenvolveu uma perspectiva cujo fundamento e critério científico é a falseabilidade e não a verificabilidade dos enunciados. Na visão popperiana, a ciência começa com problemas, não com a observação. Deparado com o problema, o cientista deve propor hipóteses para a sua resolução. Uma vez propostas, elas devem ser testadas. Caso passem nesses testes empíricos, são corroboradas e provisoriamente mantidas. Caso sejam falseadas, são abandonadas e substituídas.

Popper (1972) sugere como critério de demarcação científica a ideia de que um enunciado ou sistema de enunciados é científico se, e somente se, for falsificável empiricamente. Quanto mais falseável for um enunciado, mais arriscado e informativo ele será. Assim, por exemplo, o enunciado "chove agora em minha cidade" é muito mais arriscado que "chove agora", uma vez que a possibilidade de acontecimento ou probabilidade de ocorrência do primeiro enunciado é muito menor, em comparação ao segundo.

A proposta de Popper se contrapõe, em grande medida, à perspectiva epistemológica dos pensadores do círculo de Viena. Na postura empirista, o progresso científico ocorre a partir do acúmulo de dados. Para Popper, o progresso ocorre no processo de tentativas e erros, de conjecturas e refutações. Abandonar hipóteses incapazes de serem corroboradas significa, em alguma medida, afastar-se do falso, embora não existam garantias lógicas de que a próxima hipótese seja verdadeira e que esteja mais próxima da verdade do que da falsidade.

Na proposta do Positivismo Lógico, a atividade científica consiste na prática constante da justificação de enunciados gerais a partir da observação repetida de uma dada classe de fenômenos que se tornam base para a confirmação de leis. Para Popper, ela consiste na atividade de sugestão e teste de hipóteses para a resolução de problemas. Ambas as vertentes se preocupam, de alguma forma, em avaliar e normatizar os parâmetros de demarcação entre a ciência e a não ciência e da atividade científica. Ademais, a análise do funcionamento científico centra-se principalmente na avaliação de enunciados ou sistemas de enunciados.

Alguns pensadores como Lakatos (1989) e Kuhn (2011a) entendem que uma explicação adequada do funcionamento da ciência precisa adotar uma perspectiva historicista e estrutural, contrariando as abordagens acima citadas. Para Lakatos, a atividade científica se desenvolve a partir de programas de pesquisa. Eles são diretrizes metodológicas que norteiam as decisões referentes a construção e modificação de teorias, as quais não são consideradas elementos isolados, mas pertencentes a um dado programa.

Os programas de pesquisa possuem um núcleo rígido composto por um conjunto de hipóteses e princípios convencionalmente aceitos sistematicamente. Eles são considerados "irrefutáveis" e necessários à atividade científica no programa. Assim, por exemplo, parte do núcleo rígido do programa de pesquisa de Copérnico, destaca Lakatos (1989, p. 234, tradução nossa), é a "[...] proposição de que as estrelas constituem o sistema de referência fundamental para a física". Os cientistas que adotam um programa de pesquisa não descartam o elemento que compõem o núcleo rígido do programa, mesmo quando fatos problemáticos são constatados contra eles, como, por exemplo, refutações de teorias, através de fenômenos que os contrariem ou de previsões que não se cumprem.

Lakatos afirma que a heurística negativa do programa nos diz que certos caminhos devem ser evitados por uma teoria, proibindo que, frente a qualquer caso problemático, seja declarado falso o núcleo rígido. Desse modo, tal núcleo é preservado das refutações, em razão da existência de algumas hipóteses auxiliares, chamadas de cinturões protetores. No programa de Newton, por exemplo, havia alguns modelos do sistema solar, a teoria sobre a refração da luz na atmosfera, a forma e distribuição da massa dos planetas. Existem momentos em que os cientistas se deparam com

anomalias e refutações que levam a modificações nos cinturões protetores; a denominada heurística positiva orienta, parcialmente, as modificações que devem ser feitas no cinturão protetor do programa.

Lakatos, discordando dos positivistas lógicos e de Popper, também defende que não são teorias que estão na berlinda. A história monstra que os adeptos de uma corrente não costumam abandonar suas teorias diante do primeiro contraexemplo. De início, diante de um obstáculo, as desconfianças são sempre, por exemplo, em relação a experimentos, à precisão dos instrumentos utilizados, aos cuidados dos experimentadores. Na maioria das vezes, a história mostra que tal procedimento foi adequado.

Kuhn (2011a), também como Lakatos, adota uma postura estruturalista. A noção de paradigma aparece como um elemento central da abordagem do progresso científico deste pensador. Um paradigma estabelece e norteia a atividade de uma comunidade científica. A fase inicial de boa parte das áreas de pesquisa é marcada por um momento de luta paradigmática – chamada pré-ciência –, que se caracteriza pela existência de diversos paradigmas rivais. Quando um deles passa a ser dominante, a área de pesquisa entra num período denominado ciência normal. O paradigma pode apresentar, ao longo do tempo, anomalias sérias que podem originar uma crise científica. Ocasionalmente ocorre a emergência de novos candidatos a paradigma dominante capazes de, em tese, responder a questões e problemas que o seu paradigma rival fora incapaz de responder. Nesse momento, a área passa por um período de ciência extraordinária, de luta paradigmática. O abandono do antigo paradigma e a transição para um novo paradigma dominante marcam um momento de revolução científica na área.

A abordagem de Kuhn apresenta algumas vantagens em relação às outras propostas elencadas acima, em especial pelo seu aspecto historicista e estruturalista. Na visão de Kuhn, a atividade científica envolve muito mais do que teorias e conjuntos de enunciados passíveis de refutação. O funcionamento e o desenvolvimento das ciências envolvem elementos psicológicos, pressupostos metafísicos e metodológicos. Na visão de Kuhn, ambas as propostas, do Positivismo Lógico e de Popper, possuem uma concepção de progresso científico equivocado, dado o seu normativismo e falta de um estruturalismo. Entendemos que esses dois aspectos são elementos favoráveis a Kuhn. No tocante a Lakatos, pensamos que ele é

menos adequado do que Kuhn para os nossos propósitos, por entender que o processo de passagem de um programa de pesquisa para outro é um processo racional. Ademais, Lakatos parece não ser um historicista *stricto senso*, uma vez que prescreve uma forma ideal de se fazer ciência.

Dado estes elementos, neste livro, buscamos investigar o estatuto científico da ciência cognitiva em sua fase inicial, tomando como base a estrutura das revoluções científicas de Thomas Kuhn. Buscamos analisar, em particular, se a ciência cognitiva já começa como uma ciência normal ou se começa como pré-ciência.

A proposta de Kuhn prontifica-se, fundamentalmente, em tratar das grandes revoluções científicas. Nesta obra, abordamos o estatuto científico de uma área de pesquisa em particular. Entendemos que tal adaptação é possível e justificada, por um lado, porque nos apropriamos do arcabouço conceitual kuhniano para tratar do estatuto científico desta área. Por outro lado, o próprio Kuhn (2011a) menciona que as características das revoluções científicas podem ser igualmente examinadas através do estudo de outros episódios que não foram tão obviamente revolucionários. As equações de Maxwell, por exemplo, afetaram um grupo bem mais reduzido do que as de Einstein, porém, não foram consideradas menos revolucionárias e, por esse motivo, encontraram resistência. Kuhn (2011a, p. 74) sugere "[...] a existência de revoluções grandes e pequenas, algumas afetando apenas os estudiosos de uma subdivisão de um campo de estudos. Para tais grupos, até mesmo a descoberta de um fenômeno novo e inesperado pode ser revolucionária".

Para alcançar o nosso objetivo neste livro, o dividimos em três capítulos. No primeiro deles apresentamos a proposta de Thomas Kuhn sobre as revoluções científicas. Em um sentido geral, tratamos da concepção de paradigma, o que envolve uma espécie de modelo que estabelece e norteia a atividade de uma comunidade científica. Essa concepção de ciência determina que os paradigmas também são fundamentais para delimitar a fase em que uma área de pesquisa se encontra e para a explicação das revoluções científicas. Expomos os principais conceitos associados às diferentes fases pelas quais uma área de pesquisa pode passar ao longo de sua história, sendo elas: pré-ciência, ciência normal e progresso interno da ciência, crise e revolução científica.

No segundo capítulo expomos o surgimento da ciência cognitiva a partir do ponto de vista histórico. Apresentamos as cinco características de maior importância, as quais podem ser consideradas indicativas da atividade dessa área de pesquisa em seus primórdios: o uso de representações mentais; a crença de que os computadores eletrônicos são bons modelos explicativos de processos cognitivos; a decisão deliberada em não enfatizar certos fatores "demasiadamente complicados"; a interdisciplinaridade; e a consolidação de uma comunidade científica. Evidenciamos, ainda, alguns elementos históricos fundadores da ciência cognitiva, como os primeiros textos que fazem referência a ela, a formação dos primeiros pesquisadores, o importante papel representado pelas Conferências Macy no processo de fortalecimento de sua comunidade científica e a sua relação com um movimento intelectual anterior, chamado cibernética.

No terceiro capítulo analisamos os aspectos epistemológicos, buscando avaliar se, em seu surgimento, haveria razões suficientes para justificar a existência de uma área de pesquisa já consolidada e se ela já teria estabelecido um paradigma dominante. A fim de evidenciar as bases teóricas e metodológicas, apresentamos as suas duas principais vertentes naquele período, o cognitivismo e o conexionismo. Também denominado Inteligência Artificial (IA), o cognitivismo, em seu sentido forte e fraco, consiste na hipótese de que a cognição funciona de maneira idêntica ou semelhante aos procedimentos computacionais. Já a perspectiva conexionista, também denominada Redes Neurais Artificiais, consiste na ideia de que a cognição resulta do processamento distribuído e paralelo realizado pelas várias unidades simples que compõem uma rede neural.

Diante desses elementos históricos e epistemológicos, buscamos concluir em que medida poderíamos afirmar ou negar a existência de um paradigma dominante na formação inicial da ciência cognitiva. Em outras palavras, no arcabouço conceitual de Kuhn, teria ela começado como uma ciência normal ou teve seu início como a maioria das áreas de pesquisa, num período de disputa entre paradigmas rivais?

1

A estrutura das revoluções científicas

Apresentação

Expomos, neste capítulo, a estrutura das revoluções científicas segundo Thomas Kuhn, explicitando quatro momentos da atividade de uma área de pesquisa. Iniciamos o capítulo tratando da concepção de paradigma, um dos principais conceitos introduzidos em *A estrutura das revoluções científicas*. Os paradigmas são como uma espécie de modelo que estabelece e norteia a atividade de uma comunidade científica. Ele também é fundamental para delimitar a fase em que uma área de pesquisa se encontra e para explicação das revoluções científicas. Na segunda seção apresentamos o estágio de pré-ciência, marcado pela competição entre paradigmas rivais. Em seguida, explicitamos a concepção de ciência normal, momento em que a prática científica se normaliza em torno de um paradigma dominante. O paradigma serve de guia para a comunidade científica, empenhada em solucionar os quebra-cabeças elencados em sua agenda. Na quarta seção tratamos do período de crise, originado, dentre outros fatores, por fenômenos que fogem à explicação paradigmática vigente, denominados de anomalias. Tal período é marcado pela desconfiança no paradigma dominante e fortalecido com o surgimento de paradigmas rivais mais promissores que o atual, gerando uma revolução científica, a ser tratada na quinta seção. Parte significativa deste capítulo foi publicado em Alves e Valente (2020).

Nossa exposição se refere à clássica abordagem explicativa do funcionamento da ciência, exposto por Kuhn em *A estrutura das revoluções científicas*, divulgada pela primeira vez em 1962. Não discutiremos mudanças posteriores, em maior ou menor grau, efetivadas por esse pensador em sua perspectiva, como, por exemplo, em relação à incomensurabilidade de teorias ou à sincronicidade de paradigmas rivais, tal como descrito, por exemplo, por Hoyningen-Huene (1993). Ademais, a exposição feita aqui é ingênua e basicamente descritiva, buscando expor os pontos que nos interessam para o alcance principal de nosso objetivo neste trabalho.

1.1 A concepção de paradigma em Thomas Kuhn

A aquisição de um paradigma dominante é um dos sinais de maturidade de uma área de pesquisa. Os paradigmas correspondem a modelos ou padrões bem aceitos norteadores da atividade científica de uma comunidade. Conforme Kuhn (2011a, p. 13), "[são] realizações científicas universalmente reconhecidas que, durante algum tempo, fornecem problemas e soluções modelares para uma comunidade de praticantes de uma ciência".

Um paradigma envolve, dentre outras coisas, uma visão de mundo, princípios metodológicos, regras, teorias, comunidade científica e agenda de problemas a serem tratados pela comunidade em um dado momento.

De acordo com Kuhn (2011a, p. 220),

> [...] o termo "paradigma" é usado em dois sentidos diferentes. De um lado, indica toda a constelação de crenças, valores, técnicas etc., partilhadas pelos membros de uma comunidade determinada. De outro, denota um tipo de elemento dessa constelação: as soluções concretas de quebra-cabeças que, empregadas como modelos ou exemplos, podem substituir regras explícitas como base para a solução dos restantes quebra-cabeças da ciência normal.

Nesse contexto, por exemplo, na física, podemos citar como paradigmas as abordagens aristotélica e newtoniana, esta última tendo, na modernidade, substituído aquela. Contemporaneamente, o paradigma

da relatividade tornou-se dominante nessa área, tendo como rival, ainda não demasiadamente ameaçador, a física quântica. Na astronomia, são exemplos os sistemas geocêntrico e heliocêntrico, este tendo substituído aquele e que, apesar de algumas mudanças internas, se mantém dominante na área.

No posfácio de *A estrutura das revoluções científicas*, Kuhn (2011a) menciona que o conceito de paradigma foi examinado por Margaret Masterman, que preparou um índice analítico em que o termo havia sido utilizado, na primeira versão da referida obra, pelo menos de 22 maneiras diferentes. Após uma revisão, o autor optou por dois usos distintos do termo que poderiam englobar as demais maneiras: o emprego mais global do conceito relacionado à noção de "matriz disciplinar" e a abordagem do conceito de paradigma, em função de exemplos compartilhados. Kuhn (2011a, p. 228) alinhava o seguinte conceito:

> [...] o termo "teoria", tal como é empregado presentemente na filosofia da ciência, conota uma estrutura bem mais limitada em natureza e alcance do que a exigida aqui. Até que o termo possa ser liberado de suas implicações atuais, evitaremos confusão adotando um outro. Para os nossos propósitos atuais, sugiro "matriz disciplinar": "disciplinar" porque se refere a uma posse comum aos praticantes de uma disciplina particular; "matriz" porque é composta de elementos ordenados de várias espécies, cada um deles exigindo uma determinação mais pormenorizada.

Dentro da matriz disciplinar estão as "generalizações simbólicas", expressões utilizadas sem maiores problemas de compreensão pelos integrantes do grupo, como, por exemplo, fórmulas lógicas e matemáticas, ou termos como "os elementos combinam-se em uma proporção constante às suas massas". Na visão de Kuhn (2011a), essas expressões geralmente são aceitas pela grande maioria da comunidade. Sem elas, o grupo não teria ponto de apoio para a aplicação de técnicas de manipulação lógica e matemática. Ademais, em geral, o poder de uma ciência parece aumentar com o número de generalizações lógicas ao seu dispor.

As generalizações simbólicas funcionam em parte como leis e em parte como definições dos símbolos que elas empregam. Por exemplo, na expressão correspondente à noção de trabalho na física, em que *t* expressa

trabalho, *F* representa a força obtida a partir da multiplicação entre a massa e a aceleração e Δ_s expressa o deslocamento, podemos obter a seguinte expressão: $t = F.\ \Delta_s$. As generalizações aparecem no contexto das definições que são tautológicas e no ambiente das leis que podem, com o tempo, ser desenvolvidas e gradualmente corrigidas. O abandono de generalizações incontestáveis em uma teoria, frequentemente denominados postulados, são fatores que favorecem o agravamento de crises em uma área de pesquisa.

Outro componente da matriz disciplinar são os assim denominados "paradigmas metafísicos" ou "partes metafísicas dos paradigmas". Na física e em outras áreas como a química, eles correspondem aos compromissos coletivos, por exemplo, como as crenças de que o calor corresponde à energia cinética das partes constituintes dos corpos e que todos os fenômenos perceptivos são decorrentes da interação dos átomos. Tais componentes representam modelos que incluem uma variedade de elementos ontológicos, no que tange à realidade. Eles ajudam a estabelecer os problemas e as soluções que guiam a comunidade de um paradigma. Kuhn (2011a, p. 64) enfatiza:

> Por exemplo, depois de 1630 e especialmente após o aparecimento dos trabalhos imensamente influentes de Descartes, a maioria dos físicos começou a partir do pressuposto de que o Universo era composto por corpúsculos microscópicos e que todos os fenômenos naturais poderiam ser explicados em termos da forma, do tamanho, do movimento e da interação corpusculares. Esse conjunto de compromissos revelou possuir tanto dimensões metafísicas como metodológicas. No plano metafísico, indicava aos cientistas que espécie de entidades o Universo continha ou não continha – não havia nada além da matéria dotada de forma e em movimento. No plano metodológico, indicava como deveriam ser as leis definitivas e as explicações fundamentais: leis devem especificar o movimento e as interações corpusculares; a explicação deve reduzir qualquer fenômeno natural a uma ação corpuscular regida por essas leis.

Outro elemento fundamental dos paradigmas são os exemplares. Com essa expressão, são indicadas as soluções concretas dos inúmeros problemas ou exercícios com o quais os neófitos se deparam, desde o início de sua formação profissional, seja nos laboratórios, seja nos manuais científicos, seja ainda nos periódicos científicos. Essas soluções lhes

indicam, por meio de exemplos, as maneiras como devem proceder com seus trabalhos. Assim, como exemplifica Kuhn (2011a, p. 234), "[...] todos os físicos começam aprendendo os mesmos exemplares: problemas como o do plano inclinado, do pêndulo cônico, das órbitas de Kepler; e o uso de instrumentos como o vernier, o calorímetro e a ponte de Wheatstone".

Em geral, afirma-se que os iniciantes de qualquer área da ciência só são capazes de resolver problemas após terem aprendido sobre as teorias e algumas regras que indicam a sua aplicabilidade, inseridos em um paradigma.

Um paradigma se fortalece quando possui a capacidade de atrair um grupo duradouro de partidários, afastando-os de outros paradigmas pretendentes. Ele garante a possibilidade de resolução dos problemas científicos pelos participantes do grupo adepto aos seus pressupostos, os quais se sentem desafiados a solucioná-los.

Os praticantes de uma mesma área de pesquisa associados a um paradigma estão comprometidos com as mesmas regras, padrões e pressupostos, constituindo uma comunidade. Seus integrantes são submetidos a uma educação com uma grande extensão de literatura técnica retirada de manuais e demais revistas científicas que demarcam o limite do objeto de estudo da área, os pressupostos, métodos e metodologias científicas. Os membros da comunidade acreditam ser os responsáveis pela busca e estudo de um conjunto de objetivos em comum, que inclui o treino de seus sucessores.

Uma comunidade, em seu interior, apresenta uma ampla comunicação coesa entre seus praticantes e um julgamento profissional referente aos avanços internos, os quais tendem a ser relativamente unânimes. Kuhn (2011a, p. 223) sugere níveis de comunidades:

> A comunidade mais global é composta por todos os cientistas ligados às ciências naturais. Em um nível imediatamente inferior, os principais grupos científicos profissionais são comunidade: físicos, químicos, astrônomos, zoólogos e outros similares. Para esses agrupamentos maiores, o pertencente a uma comunidade é rapidamente estabelecido, exceto nos casos limites. Possuir a mais alta titulação, participar de sociedades profissionais, ler periódicos especializados, são geralmente condições mais do que suficientes.

Seguindo o exemplo de Kuhn (2011a), ainda existem subgrupos dentro de uma comunidade, como os biólogos adeptos da cronobiologia, que estudam a relação entre os seres vivos e o tempo, os físicos que estudam os astros e os grandes corpos, os engenheiros que apoiam o conexionismo e assim por diante. Esses tipos de comunidades são as produtoras e legitimadoras do conhecimento científico.

Os paradigmas funcionam como um mapa conceitual que permite maior profundidade e alcance explicativo da natureza. Nesse sentido, eles funcionam como um microscópio, possibilitando uma análise de um universo inacessível à visão comum. Sem a presença dos paradigmas, muitos fenômenos jamais poderiam ser explicados. Isso ocorre porque, muitas vezes, só é possível formulá-los a partir do escopo de um paradigma. Como exemplo, mencionamos a relação entre as situações-problema comuns no tratamento contemporâneo de doenças e as concepções aceitas pela comunidade científica de organismos microscópicos, as concepções éticas e profissionais sobre as células e os organismos vivos.

Kuhn se remete à presença de três diferentes enfoques que a atividade científica pode ter sobre os fatos. Em primeiro lugar, diz Kuhn (2011a, p. 46),

> [...] temos aquela classe de fatos que o paradigma mostrou ser particularmente reveladora da natureza das coisas. Ao empregá-los na resolução de problemas, o paradigma tornou-os merecedores de uma determinação mais parecida, numa variedade maior de situações.

Parte dos exemplos atrelados a essa primeira análise dos fatos implica o aperfeiçoamento e a construção de aparelhos especiais para experimentação e comprovação de teorias. Esse primeiro foco não faz menção apenas à novidade da descoberta, mas também à precisão e à segurança explicativa que o paradigma acaba por receber.

Kuhn (2011a, p. 46) alude igualmente a um segundo tipo comum de fatos a serem analisados, com base na ótica dos paradigmas:

> Uma segunda classe usual, porém mais restrita, de fatos a serem determinados diz respeito àqueles fenômenos que, embora frequentemente sem muito interesse intrínseco, podem ser diretamente comparados com as predições da teoria do paradigma.

Essa segunda classe de fatos faz referência à relação intrínseca entre os problemas teóricos dos paradigmas e os fenômenos da natureza. A aproximação teórica com o natural requer um aperfeiçoamento técnico que desafia tanto a capacidade teórica do observador quanto o treinamento e a sua imaginação. Exemplos clássicos referem-se às primeiras observações dirigidas por Galileu com seu telescópio. À época, foi extremamente dificultoso convencer a comunidade de pensadores de sua confiabilidade e adequação de uso. Tais aparelhos representam um esforço engenhoso de estreitar a relação entre teoria e natureza. No entanto, sem o treinamento ou aprofundamento ideal, podem resultar apenas em informações desconexas, ruidosas. Esse segundo tipo de trabalho com os fatos existe de uma maneira tão fundamental como o primeiro, acarretando o desenvolvimento de tecnologias e aparelhagens capazes de resolver razoavelmente o problema de comunicação entre o mundo e as teorias.

Por fim, a última e terceira classe de fatos e observações sugeridas por Kuhn (2011a, p. 48)

> Consiste no trabalho empírico empreendido para articular a teoria do paradigma, resolvendo algumas de suas ambiguidades residuais e permitindo a solução de problemas para os quais ela anteriormente só tinha chamado a atenção.

Os esforços para articular um paradigma implicam a busca de ferramentas e valores numéricos mais precisos e constantes, capazes de apresentar maior poder explicativo para maior gama de fenômenos com maior simplicidade teórica. Esse terceiro tipo de fenômenos é causado e é causador do estreitamento entre teorias e explicações dentro de um mesmo paradigma ou, pelo efeito contrário, da desvinculação de constantes, que, embora sejam muito semelhantes, por se tratar de problemas próximos, causam apenas ambiguidade. Um dos exemplos de Kuhn (2011a) concerne ao paradigma da teoria calorífica, o qual sugere o processo de aquecimento e resfriamento por meio de misturas e mudanças de estados da matéria. Segundo Kuhn (2011a), a temperatura também poderia sofrer alterações por meio de muitos outros métodos, como por combinações químicas, por fricção ou por compressão. Uma vez estabelecidos esses fenômenos sobre o aquecimento, foi necessário reformular as bases do paradigma e as suas experiências posteriores, com o intuito de elucidar os problemas até então não respondidos pelo paradigma.

Tais aspectos mostram o desenvolvimento técnico, imaginativo e interno de um paradigma. Embora esses aspectos sejam importantes, boa parte deles reflete a sofisticação do paradigma, o que Kuhn denomina de progresso interno da área de pesquisa, quando ela se caracteriza como ciência normal.

Cientistas que aderem a algum paradigma devem ter em mente a necessidade de buscar compreender certos fenômenos do mundo e ampliar o alcance explicativo do paradigma. Quando este não se encontra capaz de responder a certas questões, seus adeptos tentam refinar suas teorias, métodos, procedimentos metodológicos, instrumentos e técnicas de observação, a fim de mantê-lo sustentável.

Os paradigmas possuem um momento de origem atrelado à fundação de uma comunidade razoavelmente coesa. A fase inicial de boa parte das áreas de pesquisa é marcada por um momento carente de um paradigma dominante, anterior ao seu estabelecimento como ciência madura, denominado "pré-ciência".

1.2 Pré-ciência

A pré-ciência é a fase na qual uma área de pesquisa, uma vez inicialmente instituída, possui vários paradigmas rivais em competição igualitária. Segundo Borradori (2003), um período pré-paradigmático é conotado pelo acúmulo caótico de dados, além da pouca consolidação de uma comunidade de pesquisa que pensa ter adquirido respostas seguras sobre as entidades fundamentais que compõem o universo, das possibilidades de sua interação, além das questões que podem ser legitimamente feitas a respeito delas e quais técnicas podem ser empregadas na busca da solução de problemas a seu respeito.

Nesse momento não há a aceitação generalizada de qualquer conjunto de regras, métodos ou padrões científicos. Qualquer cientista ou grupo de cientistas precisa dispender um bom tempo criando, explicando e justificando os seus conceitos, métodos, técnicas e pressupostos, visando a se fazer entender pelos demais indivíduos ou grupos, especialmente aqueles que não aderiram à sua perspectiva.

Na ausência de um paradigma bem estabelecido ou de um candidato a paradigma assim, os fatos que se apresentam ao desenvolvimento de alguma área da atividade científica parecem ser igualmente relevantes. Somente em poucos casos, fatos com tão pouca orientação por parte de teorias preestabelecidas falam com tamanha clareza e simplicidade, para permitir o surgimento de paradigmas fundadores de uma área. Expõe Kuhn (2011a, p. 37):

> As escolas características dos primeiros estágios do desenvolvimento de uma ciência criam essa situação. Nenhuma história natural pode ser interpretada na ausência de pelo menos algum corpo implícito de crenças metodológicas e teóricas interligadas que permita seleção, avaliação e crítica. Se esse corpo de crenças já não está implícito na coleção de fatos – quando então temos à disposição mais do que "meros fatos" – precisa ser suprido externamente, talvez por uma metafísica em voga, por outra ciência ou por um acidente pessoal e histórico. Não é de admirar que nos primeiros estágios do desenvolvimento de qualquer ciência, homens diferentes confrontados com a mesma gama de fenômenos – mas em geral não com os mesmos fenômenos particulares – os descrevam e interpretem de maneiras diversas.

Em suma, enquanto pré-ciência, uma área de pesquisa carece de amadurecimento. Ela começa a amadurecer quando um dos paradigmas rivais obtém força, sendo escolhido pela maioria dos pesquisadores da área, convergindo nas atividades e atenções. Um paradigma prevalece por parecer ser melhor que os seus rivais, prometendo, em tese, possuir maior poder explicativo e preditivo que os demais.

A aceitação de um paradigma envolve também a aceitação de suas teorias. Pode acontecer de teorias rivalizarem dentro de um mesmo paradigma na tentativa de explicar certos fenômenos. Nesse caso, não é o paradigma o alvo de escolha, mas suas teorias. No entanto, quando teorias estão associadas a paradigmas distintos, a escolha de uma delas também direciona a escolha do paradigma a que ela está submetida.

Embora não exista um critério de qualidade rigoroso para a aceitação de teorias, Kuhn (2011b, p. 341) explicita cinco características

indicadoras de qualidade extraídas de suas investigações ao longo da história do desenvolvimento da ciência:

> Primeiro, uma teoria deve se conformar com precisão à experiência: em seu domínio, as consequências dedutíveis da teoria devem estar em clara concordância com os resultados da experimentação e da observação existente. Segundo, uma teoria deve ser consistente, não apenas internamente ou auto consistente, mas também com outras teorias correntes aplicáveis a aspectos da natureza que são afins. Terceiro, ela deve ter uma extensa abrangência; em particular, as consequências da teoria devem ir muito além das observações, leis ou subteorias particulares cuja explicação motivou sua formulação. Quarto, e fortemente relacionado, ela deve ser simples, levando ordem a fenômenos que, em sua ausência, permaneceriam individualmente isolados e coletivamente confusos. Quinto – um item um pouco incomum, mas de importância crucial para as decisões científicas efetivas –, uma teoria deve ser fértil em novos achados de pesquisa, deve abrir portas para novos fenômenos ou a relações antes ignoradas entre fenômenos já conhecidos.

Kuhn (2011a) explicita dois tipos de valor que podem servir de base para a escolha e a qualificação de teorias e paradigmas: cognitivos e extracognitivos. Segundo Laudan (1984, p. xii, tradução nossa): "[os valores cognitivos] representam uma propriedade de teorias que supomos serem constitutivas de uma 'boa ciência'". Para o caso das teorias, os principais valores cognitivos são adequação empírica, consistência, poder explicativo, simplicidade e fecundidade.

A adequação empírica supõe que a teoria se ajuste aos dados disponíveis, mostrando-se capaz de poderes preditivos sobre esses dados e de relatar algo a respeito de fenômenos. Visto que a ciência encontra suas justificativas no âmbito físico através da observação empírica, é essencial que uma teoria trate de problemas do mundo empírico.

A tese da consistência pressupõe que uma teoria não pode se firmar por meio de pressupostos que afirmem e neguem, ao mesmo tempo, sob o mesmo aspecto, uma mesma característica sobre fenômenos. Teorias incapazes de corresponder a esse requisito são inconsistentes ou contraditórias e, assim, não seriam informativas a respeito do mundo.

Ademais, uma boa teoria necessita da capacidade de inserir-se nos meandros da comunidade científica, tanto quanto na própria linguagem, facilitando sua compreensão pelos demais membros da comunidade; independentemente de suas propostas, as terminologias devem ser aceitas e compreendidas pela comunidade como um todo.

O poder explicativo expressa que as teorias sejam destinadas a analisar os principais fenômenos do mundo dentro de seu escopo. Elas precisam definir na mais ampla extensão e profundidade as leis que regem os processos e as estruturas da natureza.

A simplicidade acrescenta que uma boa teoria deve ser simples e clara, precisando ser conceitualmente capaz de ser formalizada e inteligível. Nesse sentido, entende-se que um dos objetivos da ciência é transformar o confuso no claro, com simplicidade. A ciência, em tese, é capaz de desvelar leis que garantem e regem os aspectos mais gerais do mundo. Caso uma teoria não seja capaz de apresentar alguma previsibilidade a propósito do mundo, devido à aparição de inúmeras anomalias que a refutam, é necessário sua reformulação ou surgimento de uma nova teoria, capaz de explicar os eventos que as outras são incapazes de fazer.

Segundo Kuhn (2011a), a fecundidade é o mais importante dos valores cognitivos. É necessário à teoria ser fecunda para o conhecimento do mundo, capaz de exibir novos fenômenos ou novas implicações que garantam a possibilidade de se observar um mesmo fenômeno de maneira diferente. Uma teoria incapaz de apresentar contribuições às questões referentes ao mundo ou a questões que ela pretende tratar nada ou pouco oferece ao desenvolvimento da área de pesquisa. Ademais, pressupõe-se que toda boa teoria seja fecunda a novos problemas, fornecendo quebra-cabeças aos pesquisadores da área, motivando-os e desafiando-os a montá-los, seguindo as regras do paradigma em que estão inseridos.

Os valores cognitivos, conforme expressos, não se garantem axiomaticamente como critérios de qualidade de teorias. Contudo, são idealizações identificadas na grande maioria das teorias científicas consideradas de qualidade. Embora tais valores tenham certo rigor, eles não podem servir como critério objetivo e definitivo para a escolha entre teorias ou paradigmas, dada certa imprecisão e conflito entre essas características. Há situações, por exemplo, nas quais a simplicidade pode ditar uma

escolha. Em outros casos, a relevância pode ser atribuída à consistência. Conforme Kuhn (2011b, p. 343):

> A Astronomia heliocêntrica, que exigia o movimento da Terra, era inconsistente com as explicações científicas disponíveis desses e de outros fenômenos terrestres. Assim, tomado de maneira isolada, o critério da consistência fala de maneira inequívoca a favor da tradição geocêntrica.
>
> A simplicidade, no entanto, favorecia Copérnico, mas somente quando avaliada de modo específico [...] se examinássemos a quantidade de expediente matemático exigido para explicar não os movimentos quantitativos detalhados dos planetas, mas seus aspectos qualitativos gerais (elongação limitada, movimento retrógado e afins), veríamos, como bem sabe qualquer criança de escola, que Copérnico requer apenas uma circunferência por planeta e Ptolomeu, duas.

Assim, dois cientistas podem chegar a conclusões e escolhas diferentes sobre qual o melhor paradigma, pois esses critérios podem ser interpretados de maneira diversa: talvez os cientistas concordem nos critérios, mas discordem em suas valorações. Conforme exemplifica Kuhn (2011a), aquilo que, para Einstein, era considerado demasiado complexo ou incongruente, no que tange à teoria dos *quanta*, na qual se declara que a emissão e a absorção de energia eletromagnética dos corpos ocorrem através de "pacotes", ao contrário do que é sustentado pela teoria ondulatória clássica, poderia ser, para Bohr, apenas uma dificuldade da época, a ser superada. Frisa Kuhn (2011b, p. 344):

> [...] nenhuma lista de critérios já proposta é de fato útil [...] devemos levar em conta características que variam de cientista a cientista, sem comprometer com isso sua adesão aos cânones que tornaram a ciência científica. Embora existam e possam ser descobertos (sem dúvida, os critérios de escolha com que comecei este artigo estão entre eles), esses cânones não são, por si sós, suficientes para determinar as decisões de cada cientista. Para isso, os cânones compartilhados teriam de ser elaborados de modo que variem de um indivíduo a outro.

Os critérios de escolha levam em consideração os elementos resultantes das experiências anteriores de cada cientista até o momento em que foi necessário efetuar uma escolha. Dentro das experiências está o campo em que o indivíduo atuou, quão bem-sucedido ele foi, quanto tempo ele trabalhou na área e a quantidade de conceitos e técnicas contestados pelo novo paradigma.

Além dos valores cognitivos, para Kuhn (2011a), há os extracognitivos, que interferem no progresso e na consolidação da ação científica. Tais valores são relevantes nos momentos de aceitação dos paradigmas e de escolha de suas regras metodológicas. Segundo Lacey (1998), os valores extracognitivos[1] dizem respeito a crenças, deliberações, fins, desejos e outros estados intencionais. Eles estão intrinsicamente relacionados com instituições, ecossistemas e situações sócio-históricas. Alves (2013, p. 196) observa:

> Fatores psicológicos, como a crença, sentimento ou esperança no poder explicativo de uma teoria, a intuição de pesquisadores a respeito de uma matriz disciplinar, a possibilidade de adequação do paradigma com possíveis valores ou princípios morais, sociais, podem auxiliar na escolha de um determinado paradigma em detrimento de outro. No entanto, apesar da influência de valores extra-cognitivos como estes serem relevantes, a escolha de um paradigma está amparada em valores cognitivos, os desideratos das teorias científicas que o constituem.

A elaboração dos valores cognitivos ocorre, também, influenciada pelas valorações constituídas pela comunidade científica. Muitas vezes garantidas pela existência de valores pessoais e sociais, elas interferem para que a comunidade continue coesa. A forma como um paradigma é escolhido, em vez de outro, depende não só necessariamente dos valores cognitivos, mas também da ação valorativa individual e coletiva. Visões de mundo, concepções religiosas, políticas, econômicas e sociais podem influenciar na composição e escolha de paradigmas, embora não sejam eles os sustentáculos da manutenção do paradigma.

[1] Lacey não utiliza o conceito "extracognitivo". Para designar estes tipos de valores em seu arcabouço conceitual, ele faz uso dos termos "valores" ou "valores sociais".

Imaginemos um caso em que haveria duas teorias, ambas inconsistentes entre si, mas que se fundamentassem nos mesmos dados empíricos, de tal modo que os dados não forneceriam uma base para decidir entre ambas as teorias. O critério para escolher entre uma ou outra pode ser constituído por valores sociais ou morais, que possuiriam um poder de escolha capaz de decidir quais linhas de pesquisas seriam priorizadas.

Tal maleabilidade permite a pensadores como Lakatos (1979, p. 178) asseverar que os elementos kuhnianos de escolha entre teorias e paradigmas são basicamente subjetivos e uma "[...] questão psicológica de massas". Kuhn (2011b) rebate a crítica ao argumentar que a escolha de paradigma depende de uma mescla de fatores objetivos, como a precisão, e fatores subjetivos, ou critérios compartilhados e individuais.

Comumente, a luta entre paradigmas, embora possa levar um grande tempo e esforço, tende a chegar a um vencedor. É natural que um paradigma acabe sendo mais aceito pela maioria dos pesquisadores de uma área, constituindo uma comunidade científica. Nesse ponto, a área entra em um novo momento chamado ciência normal.

1.3 Ciência normal e progresso interno ou paradigmático

O comprometimento de uma comunidade com um único paradigma é pré-requisito para o que Kuhn (2011a) chama de "ciência normal". Borradori (2003, p. 210) descreve a passagem do período de pré-ciência para ciência normal como:

> [...] a prática científica se normaliza em torno da instituição de um "paradigma", que representa uma mescla normativa de teoria e de método. Uma amálgama, no qual se juntam um espectro de postulados teóricos, uma determinada visão de mundo, dos modos de transmissão dos conteúdos da ciência, além de uma série de técnicas de pesquisa.

A predominância de um paradigma propicia a criação de publicações especializadas, a fundação de sociedades de especialistas e a reivindicação de valorização desses elementos nos currículos. Ressalta Kuhn (2011a, p. 40):

> Quando um cientista pode considerar um paradigma como certo, não tem mais necessidade, nos seus trabalhos mais importantes, de tentar construir seu campo de estudos começando pelos primeiros princípios e justificando o uso de cada conceito introduzido.

Em princípio, no período de pré-ciência, um paradigma não passa de uma promessa de explicação para certos fenômenos do mundo. Na ciência normal, tem-se a atualização dessas promessas para a comunidade, aplicando-se as predições do paradigma aos conhecimentos e fatos que o paradigma julga relevantes, aumentando a sua capacidade explicativa e refinando o próprio paradigma, através do que Kuhn (2011a, p. 44) identifica como progresso interno:

> Poucos dos que não trabalham realmente com uma ciência amadurecida dão-se conta de quanto trabalho de limpeza desse tipo resta por fazer depois do estabelecimento do paradigma ou de quão fascinante é a execução desse trabalho. Esses pontos precisam ser bem compreendidos. A maioria dos cientistas, durante toda a sua carreira, ocupa-se com operações de acabamento. Elas constituem o que chamo de ciência normal. Examinado de perto, seja historicamente, seja no laboratório contemporâneo, esse empreendimento parece ser uma tentativa de forçar a natureza a encaixar-se dentro dos limites preestabelecidos e relativamente inflexíveis fornecidos pelo paradigma. A ciência normal não tem como objetivo trazer à tona novas espécies de fenômenos; na verdade, aqueles que não se ajustam aos limites do paradigma frequentemente nem são vistos. Os cientistas também não estão constantemente procurando inventar novas teorias; frequentemente mostram-se intolerantes com aquelas inventadas por outro.

O progresso interno seria aquele que tem como resultado uma maior especialização e aperfeiçoamento dos métodos, modelos e princípios de uma área de pesquisa. Os adeptos de uma ciência normal trabalham com as teorias e fenômenos fornecidos e explicados pelo paradigma vigente. Essas características remetem ao conceito de especialização. Alves (2013, p. 199) sublinha que os adeptos de seus respectivos paradigmas costumam agir da seguinte maneira:

> Os participantes de um paradigma costumam defendê-lo com todo vigor possível. Existe uma crença quase que incontestável no poder explicativo e preditivo das teorias que constituem um paradigma, na sua adequação e correção, ou na possibilidade de seu aprimoramento.

Há, ainda, um quadro típico de problemas e expectativas a serem resolvidos pelos praticantes e adeptos de um paradigma. Os chamados "quebra-cabeças" são problemas estipulados pelo poder explicativo do paradigma, embora, em geral, seus resultados não visem a uma novidade inesperada. Quebra-cabeça, para Kuhn (2011a, p. 59), "[...] indica, no sentido corriqueiro em que empregamos o termo, aquela categoria particular de problemas que servem para testar nossa engenhosidade ou habilidade na resolução de problemas".

Ao adotar um paradigma, a comunidade científica também adquire uma série de quebra-cabeças a serem trabalhados. Em certa medida, a comunidade tende a aceitar apenas esses quebra-cabeças como problemas genuínos a serem resolvidos pelos seus integrantes. É comum que outros quebra-cabeças, os aceitos anteriormente por outro paradigma, passem a ser rejeitados pela comunidade, por se tratar, pela perspectiva do paradigma vigente, de metafísica, ou por serem pertinentes às especialidades de outra área. Eles também podem ser rejeitados por serem demasiadamente problemáticos ou por necessitarem de um gasto exagerado de energia em um dado momento. Kuhn (2011a) destaca que a ciência contemporânea tende a se desenvolver tão rapidamente devido à natureza de seus problemas, dado que seus praticantes, na concepção do paradigma, procuram se concentrar em questões que podem ser resolvidas apenas com os pressupostos fornecidos e graças à sua engenhosidade.

O empreendimento científico de uma área enquanto ciência normal, em seu conjunto, motiva e desafia seus adeptos a encaixar as peças de seus quebra-cabeças. Não obstante, um quebra-cabeças não consiste apenas em um problema solúvel no escopo do paradigma. É necessário que ele satisfaça certas regras que limitam tanto as possibilidades de sua solução aceitável quanto os possíveis métodos necessários e instrumentos preferíveis para encontrá-la. Tais regras valem para os problemas teóricos e para os problemas práticos. Elas criam uma relação de compromisso

do pesquisador para com o paradigma, sendo preciso respeitar os seus pressupostos adotados. Enquanto estes forem admitidos, auxiliam nas resoluções dos quebra-cabeças e permitem que os demais membros da comunidade as aceitem e possam passar para outros problemas. Segundo Mendonça e Videira (2007, p. 171):

> A rigor, a ciência normal calcada no paradigma instaura o consenso nos três principais níveis da pesquisa. De fato, sob a égide de um paradigma, uma certa comunidade científica não discute mais sobre quais fatos devem ser investigados, quais métodos a serem empregados e o que se aceita como soluções. Em suma, ao praticar a ciência normal, os pesquisadores lidam – de maneira homogênea – com as questões ontológicas, metodológicas e epistemológicas.

Uma das críticas feitas à abordagem de Kuhn (2011a) assinala que seu conceito de ciência normal é pouco natural, ironicamente, não histórico, uma vez que, mesmo nas comunidades mais coesas, não existe tamanha padronização, com ausência de desacordos. No entanto, embora descreva esse momento como possuindo tal identidade coletiva, o autor não descarta a possibilidade de desacordos pontuais encontrados, por exemplo, na proposta momentânea de teorias diferentes para a explicação de certos fenômenos, em particular de anomalias.

Um caso ilustrativo dessas discordâncias internas é a existência de teorias atualmente concorrentes na física, mais propriamente, na cosmologia, a respeito da explicação de como é feita a matéria e como ela se comporta no nível subatômico. Conforme explicam Corbett *et al.* (2012), com a descoberta da partícula do bóson de Higgs, que dá massa às partículas elementares, a teoria com mais adeptos, denominada "modelo padrão", deu um passo significativo na montagem do quebra-cabeças em questão, levando em consideração explicações a respeito da matéria escura, que serviria como uma espécie de cola do cosmos, dando liga ao universo, mantendo-o estável. Entretanto, tal teoria deixa uma série de lapsos explicativos, em especial na explicação da matéria escura, composta por partículas que o modelo padrão falha em explicar. Isso permite a emergência de outras teorias, como as de "Higgs composto", "dimensões extras" e "supersimetria". Inúmeros testes e experimentos empíricos são

realizados pela comunidade toda para corroborar ou mesmo provar as hipóteses e afirmações de cada teoria.

Apesar das diferenças teóricas, os pesquisadores inclinados sobre o quebra-cabeça nessa área da física seguem e são adeptos do paradigma da Teoria da Relatividade. É esta adesão que, na perspectiva de Kuhn (2011a), possibilitará a resolução do quebra-cabeças, se os pesquisadores se mantiverem entusiasmados e desafiados pela sua resolução. Mesmo havendo uma unidade entre os pesquisadores nessa área de pesquisa, há um grupo de dissidentes, defensores da Mecânica Quântica, o qual se mantém avesso a essa matriz disciplinar. Ele ainda é minoritário e, de alguma forma, permanece um tanto afastado do grupo geral, dos eventos, das publicações, até dos financiamentos de pesquisa. É possível, uma vez que as anomalias do paradigma vigente aumentem significativamente, que esse grupo venha a crescer, favorecendo um momento de crise paradigmática e gerando uma revolução científica que envolveria mudanças na física e, provavelmente, no corpo científico como um todo.

Enquanto ciência normal, há uma coesão comunitária na área, o que favorece ou mesmo possibilita o seu desenvolvimento e, por sua vez, de um paradigma. Esse progresso interno consiste no aprimoramento, dentre outras coisas, de melhores técnicas para a resolução de quebra-cabeças. Nas ciências médicas, por exemplo, o empenho de sua comunidade auxiliou no desenvolvimento do conhecimento, das intervenções e da cura do câncer, através de diferentes técnicas como cirurgia, quimioterapia, radioterapia e transplante de medula óssea.

Uma vez que os pesquisadores não precisam despender tempo para as questões de cunho filosófico ou metodológico, não é mais necessário discutir quais fatos devem ser levados em consideração, quais quebra-cabeças devem ser resolvidos nem quais soluções devem ser encontradas. Com esse consenso, o progresso interno empreendido por uma comunidade científica tende ao crescimento contínuo e cumulativo. Nesse sentido, falar de progresso científico durante a fase de ciência normal é sinônimo de progresso interno. Conforme explica Mendonça (2012, p. 539):

> O progresso científico, no sentido de aprofundamento no conhecimento dos fatos, de aperfeiçoamento dos métodos de investigação e avanço nos resultados esperados é uma decorrência do consenso engendrado pelo paradigma. Nesse sentido, pode-se

> afirmar que, durante a ciência normal, o conhecimento progride de forma linear e cumulativa. Esse aspecto da descrição kuhniana não é tão frequentemente discutido; às vezes, ele não é sequer notado, o que não ocorre com a sua concepção descontinuísta de progresso concernente às revoluções científicas.

Assim, Kuhn (2011b) apresenta uma imagem sobre progresso científico em duas direções distintas, mas complementares: o paradigmático e o revolucionário. Isso o levou à sua famosa tese de que o desenvolvimento de qualquer área da ciência ocorre através de uma profunda e sutil tensão essencial entre os limites da normalidade e da revolução.

Apesar do período da ciência normal ser bem-sucedido em seu empreendimento, um paradigma, após sua consolidação, pode apresentar muitas limitações, quer naquilo que tange à sua precisão, quer no seu aprimoramento interno. Pode apresentar, ao longo do tempo, anomalias sérias, o que pode resultar em uma crise científica. A aparente solidez do paradigma começa a desmantelar e o período de ciência normal se transforma em um momento de rupturas. Nesse estágio, são questionados seus métodos, técnicas e fundamentos. A história da ciência é marcada por vários exemplos desse momento. A passagem do sistema geocêntrico para o heliocêntrico, na astronomia, é um deles. Investigamos, na próxima unidade, os elementos que levam um paradigma a entrar em crise e quais são as possíveis soluções para essa situação.

1.4 Crise

A atividade da ciência normal consiste em um empreendimento coletivo voltado para a resolução de quebra-cabeças estabelecidos na agenda do paradigma. Entretanto, é comum, nesse momento, que fenômenos anômalos sejam apreendidos, exigindo, por vezes, uma capacidade explicativa para além do paradigma dominante para tais fenômenos.

Kuhn (2011a, p. 77) assinala que a "[...] ciência normal não se propõe descobrir novidades no terreno dos fatos ou teorias; quando é bem-sucedida, não as encontra". As chamadas "anomalias" são reconhecidas a partir do momento em que as expectativas do paradigma vigente são suplantadas por fenômenos que ou são extremamente difíceis de serem

explicados ou são tão anômalos que é impossível explicá-los a partir das teorias do paradigma. No primeiro caso, é comum que a teoria tente se ajustar, a fim de assimilar o novo fenômeno ao seu escopo explicativo, transformando o inesperado em esperado. O segundo caso é caracterizado pela necessidade da criação de novos conceitos e de um novo vocabulário para analisar o fenômeno, uma vez que a observação e a assimilação do novo fato a uma teoria estão muitas vezes associadas à descoberta do fenômeno.

O surgimento de anomalias pode ser resultado da descoberta de algum elemento natural novo, como, por exemplo, algum elemento químico ou o comportamento de algum elemento já conhecido. O enfrentamento de anomalias pode significar a reformulação conceitual em algum nível que, por sua vez, pode exigir alguma modificação no paradigma vigente. Kuhn (2011a, p. 82) alude ao exemplo de Lavoisier para explicar a relação entre o surgimento de anomalias, a descoberta de novos elementos e a criação de paradigmas com maior poder explicativo:

> O que Lavoisier anunciou em seus trabalhos posteriormente a 1777 não foi tanto a descoberta do oxigênio como a teoria da combustão pelo oxigênio. Essa teoria foi a pedra angular de uma reformulação tão ampla da química que veio a ser chamada de revolução química [...] Muito antes de desempenhar qualquer papel na descoberta de um novo gás, Lavoisier convenceu-se de que havia algo errado com a teoria flogística [...] O trabalho sobre o oxigênio deu forma e estrutura mais precisas à impressão anterior de Lavoisier de que havia algo errado na teoria química corrente.

A experiência prévia de se observar a dificuldade de analisar um fenômeno a partir de um paradigma é um indicativo importante da necessidade de sua revisão. Muitas vezes, a percepção de uma anomalia, ou seja, de um fenômeno que o paradigma não está preparado para explicar, desempenha um papel importante para o ambiente das descobertas científicas. Apesar de não serem os únicos indicativos da necessidade da criação de paradigmas, as descobertas incluem uma série de elementos impactantes nessa mudança, como, por exemplo, a consciência prévia da existência de anomalias, o reconhecimento do plano conceitual e metodológico do fenômeno e, consequentemente, mudanças no paradigma

que, com frequência, estão relacionados com uma forte resistência oriunda da comunidade de pesquisadores mais ortodoxos.

Habitualmente, as descobertas emergem com muita dificuldade, seguindo um caminho extremamente contraintuitivo ou aparadigmático. Em princípio, as observações de cientistas estão atreladas à previsibilidade do paradigma a que são adeptos, mesmo em situações que mais tarde podem ser entendidas como anomalias. Enfatiza Kuhn (2011a, p. 91):

> Contudo, uma maior familiaridade dá origem à consciência de uma anomalia ou permite relacionar o fato a algo que anteriormente não ocorreu conforme o previsto. Essa consciência da anomalia inaugura um período no qual as categorias conceituais são adaptadas até que o que inicialmente era considerado anômalo se converta no previsto. Nesse momento completa-se a descoberta.

Em sua fase de construção, espera-se que um paradigma consiga, com alguma eficiência, explicar boa parte dos fenômenos e objetos pretendidos pela sua área de atuação. Com o passar do tempo, quando o paradigma se consolida e a área de pesquisa passa ao estatuto de ciência normal, pode ocorrer o desenvolvimento interno do paradigma, com um refinamento metodológico e conceitual que, em geral, leva a comunidade científica à especialização de sua visão de mundo e a um aumento na precisão entre os fenômenos e as teorias. Kuhn (2011a, p. 91) salienta: "A ciência torna-se sempre mais rígida." Por maior que sejam a precisão e o alcance explicativo de um paradigma, sempre há a possibilidade de haver anomalias sérias e algum indicativo da necessidade de mudança dele. A sua comunidade, no entanto, dificilmente o abandonará, sem a existência de provas consistentes, em vários laboratórios, circunstâncias ou situações de sua insuficiência e das impossibilidades de reformulá-lo.

Quando a existência de anomalias atinge certo nível, seja em sua quantidade, seja em sua relevância, a comunidade científica se encontra em uma situação em que é necessária a assimilação desses fenômenos, quer a partir da reformulação ou aprimoramento do paradigma, quer através da criação de um novo paradigma, capaz de explicar os fenômenos novos e os já conhecidos. Ao longo dessa mudança, é preciso que algumas crenças aceitas no paradigma vigente sejam descartadas e substituídas por outras.

Em boa parte dos casos, a presença de anomalias persistentes ocasiona, na comunidade científica, um período de crise crescente. As chamadas "crises" são decorrentes da insegurança gerada em face do fracasso constante da ciência normal em produzir previsões e explicações adequadas aos fenômenos da natureza. Frisa Kuhn (2011a, p. 95): "O fracasso das regras existentes é o prelúdio para uma busca de novas regras".

Os cientistas, mesmo já com pouquíssima confiança nas previsões e explicações do paradigma dominante, com frequência não são capazes de se desvencilhar do veículo de sua crise, sem uma opção razoavelmente segura e com poder explicativo potencialmente superior ao seu anterior. Baseando-se em fatos históricos, como a Revolução Copernicana, por exemplo, uma crise é superada e um paradigma é abandonado apenas quando existir alguma alternativa disponível potencialmente capaz de substituir o vigente.

Na visão de Kuhn (2011b), os motivos que levam cientistas a abandonar um paradigma englobam muito mais do que a comparação entre teorias, envolvendo a comparação mútua entre os paradigmas com a natureza, juntamente com a aceitação da comunidade científica. A resistência dos membros da comunidade, às vezes, é tamanha em tentar salvar o paradigma das crises que é comum observarmos uma constante tentativa de modificações na matriz e proposição de hipóteses *ad hoc.* Enfatiza Kuhn (2011a, p.109):

> Tal como artistas, os cientistas criadores precisam, em determinadas ocasiões, ser capazes de viver em um mundo desordenado – descrevi em outro trabalho essa necessidade como "a tensão essencial" implícita na pesquisa científica.

O momento de crise não é demarcado pelo abandono absoluto de todo e qualquer paradigma. Diferentemente da pré-ciência, a crise é caracterizada pelo momento de perda de credibilidade do paradigma dominante e surgimento de novos paradigmas rivais promissores, diferentes e contraditórios com o anterior, os quais competem pela aceitação da comunidade científica.

Kuhn (2011b) aborda a relação tênue entre os quebra-cabeças da ciência normal e as anomalias constantes nos momentos de crise.

Eles também podem ser vistos como anomalias e um indicativo de crise. Embora não exista uma linha divisória precisa entre a relação das anomalias e os quebra-cabeças, a crise tende a fomentar o enfraquecimento das suas resoluções. Nesse sentido, os paradigmas vigentes, através de suas teorias, podem ter duas estratégias de atuação, em relação às anomalias: ou elas partem de um eterno confronto com as anomalias ou jamais as confrontam diretamente. Contudo, em ambos os casos, parece sempre existir algum indicativo de problema no paradigma.

As anomalias possuem forte impacto na geração de uma crise. Provavelmente não há uma resposta direta para os processos que levam anomalias a criar uma crise em uma área de pesquisa, dada a existência de muitos fatores externos e internos extremamente impactantes para tanto. Kuhn (2011a, p.113) descreve alguns exemplos de anomalias que resultaram no desenvolvimento de crises:

> Algumas vezes uma anomalia colocará claramente em questão as generalizações explícitas e fundamentais do paradigma – tal como o problema da resistência do éter com relação aos que aceitaram a teoria de Maxwell. Ou, como no caso da revolução copernicana, uma anomalia sem importância fundamental aparente pode provocar uma crise, caso as aplicações que ela inibe possuam uma importância prática especial – neste exemplo para a elaboração do calendário e para a astrologia.

Anomalias surgem com certa frequência na atividade de um paradigma. Algumas parecem ser algo muito mais além do que um possível novo quebra-cabeças da ciência normal. Embora cientistas e investigadores tentem resolver os problemas que são cada vez mais constantes, ao passo que mais adaptações são inseridas, para esses investigadores a disciplina jamais será a mesma. Na crise, mesmo com um paradigma dominante, a comunidade já não confia mais nele como antes. Mesmo as soluções anteriormente bem aceitas passam a ser alvo de questionamentos, de desconfianças.

Esse período de confusão é marcado pela ausência de coesão entre o corpo do paradigma e a constante assimilação de hipóteses *ad hoc*. As crises iniciam-se com a falta de credibilidade do paradigma dominante

e, consequentemente, com o esvaziamento das regras orientadoras da pesquisa, ao longo do período de ciência normal.

Conforme o ditado popular, depois da tempestade vem a bonança. Em geral, depois da crise vem a revolução, de que tratamos a seguir. Explicitamos seu impacto nas comunidades científicas, incluindo as mudanças das visões de mundo e o caráter necessário das revoluções para o progresso científico.

1.5 Revoluções científicas

Uma crise pode terminar de três maneiras. A primeira possibilidade consiste na revelação de que o próprio paradigma posto em xeque seja capaz de resolver os motivos da crise. A segunda opção decorre da resistência do problema. Nesse caso, tanto as abordagens novas quanto as clássicas se revelam insuficientes. O problema é considerado insolúvel para o momento e é engavetado provisoriamente, para que possa ser resolvido por futuras gerações, as quais podem dispor de métodos e instrumentos mais sofisticados e adequados para o seu tratamento. A terceira opção acaba com o surgimento de um novo candidato a paradigma dominante e com o confronto entre ele e o vigente. Este último modo de resolução propicia a entrada da área de pesquisa num período denominado "ciência extraordinária". O termo "ciência extraordinária" faz referência a um momento de grandes transições que são fundamentais para o desenvolvimento da atividade científica, gerando novos conhecimentos sobre a realidade.

Na transição de um paradigma para o outro, em períodos de crise, ocorre o surgimento de uma nova tradição de ciência normal. A despeito de ter que ser capaz de responder, senão a todas, pelo menos a uma parte considerável das questões que o paradigma anterior era capaz de responder e que ainda se mantêm significativas, e mais às novas questões postas, a nova matriz jamais será proveniente da acumulação de dados obtidos através de sua antecessora. O novo paradigma é, antes de tudo, uma reformulação da área, a partir de novos princípios e pressupostos. Muitas vezes implica mudanças drásticas nos métodos e práticas anteriormente aceitos, o que acarreta, frequentemente, elementos contraditórios à perspectiva anterior.

Completada a transição paradigmática, os cientistas modificam, dentre outras coisas, as suas concepções de mundo e tomam os mesmos dados e observações realizados anteriormente, em função de um novo sistema de relações, averiguando-os a partir de um novo quadro de conceitos.

Em geral, os novos paradigmas surgem, ainda que em fases iniciais, antes que a crise tenha encontrado o seu apogeu. Esclarece Kuhn (2011a, p. 118):

> Confrontado com uma anomalia reconhecidamente fundamental, o primeiro esforço teórico do cientista será, com frequência, isolá-la com maior precisão e dar-lhe uma estrutura. Embora consciente de que as regras da ciência normal não podem estar totalmente certas, procurará aplicá-las mais vigorosamente do que nunca, buscando descobrir precisamente onde e até que ponto elas podem ser empregadas eficazmente na área de dificuldade, de torná-la mais nítida e talvez mais sugestiva do que era ao ser apresentada em experiências cujo resultado pensava-se conhecer de antemão [...] dado que nenhuma experiência pode ser concebida sem o apoio de alguma espécie de teoria, o cientista em crise tentará constantemente gerar teorias especulativas que, se bem-sucedidas, possam abrir o caminho para um novo paradigma e, se mal-sucedidas, possam ser abandonadas com relativa facilidade.

O desenvolvimento desses procedimentos, que exigem uma força criativa extraordinária, ocorre, por vezes, ao concentrar o enfoque científico em determinada área para o reconhecimento de anomalias. Essas observações, em momentos de crise, são fundamentais para o desenvolvimento da atividade científica. Não raro, a forma do novo paradigma é decorrente desse movimento extraordinário, sendo necessário, após a superação da crise, migrar para um novo período de ciência normal, o qual exige um refinamento conceitual.

Em linhas gerais, quando um grupo de pesquisadores desenvolve um novo paradigma que atrai a maior parte dos praticantes da área de pesquisa contemporânea e as novas gerações, é comum que as escolas mais antigas comecem a desaparecer gradualmente. Kuhn (2011a) observa ser comum o desaparecimento dessas escolas, com o surgimento de novos

paradigmas, mas sempre existem alguns indivíduos com forte elo com as concepções mais antigas. Eles são habitualmente excluídos ou forçados a unir-se a um novo grupo ou acabam se isolando da comunidade.

A transição para um novo paradigma caracteriza, do ponto de vista histórico, uma revolução científica. Revoluções científicas e transições da pré-ciência para a ciência normal constituem momentos de mudança paradigmática. Kuhn (2011a, p. 125) considera "[...] revoluções científicas aqueles episódios de desenvolvimento não-cumulativo, nos quais um paradigma mais antigo é total ou parcialmente substituído por um novo, incompatível com o anterior".

Embora nem sempre seja possível mensurar paradigmas, espera-se que o novo paradigma seja capaz de resolver muitos problemas que o paradigma anterior não conseguia; que os seus princípios metodológicos sejam mais adequados para a solução de problemas; que sejam feitas previsões e explicações mais precisas de determinados fenômenos; que seu arcabouço conceitual seja mais rigoroso e adequado aos elementos envolvidos nas investigações. Estes e outros elementos indicam que o novo paradigma é capaz de explicar, de modo mais adequado, determinado conjunto de fenômenos no mundo. São tais fatores, dentre outros, originados pela troca de um paradigma dominante por outro, que caracterizam o progresso na ciência para Kuhn.

A mudança de paradigma é marcada pela subdivisão dos membros da comunidade científica e a crescente perda de credibilidade do paradigma vigente. O sentimento de funcionamento falho ou defeituoso é uma constante que permeia o ambiente dos pesquisadores, que deixam de acreditar que o paradigma existente é capaz de explorar adequadamente o funcionamento do universo dos fenômenos a que se dispõe explicar. Kuhn (2011a, p. 126, grifo do autor) faz um paralelismo entre as revoluções políticas e as revoluções científicas:

> As revoluções políticas visam realizar mudanças nas instituições políticas, mudanças essas proibidas por essas mesmas instituições que se quer mudar. Consequentemente, seu êxito requer o abandono parcial de um conjunto de instituições em favor de outro. E, nesse ínterim, a sociedade não é integralmente governada por nenhuma instituição. De início, é somente a crise que atenua o papel das instituições políticas, do mesmo modo que atenua o papel dos

> paradigmas [...] A essa altura, a sociedade está dividida em campos ou partidos em competição, um deles procurando defender a velha constelação institucional, o outro tentando estabelecer uma nova. Quando ocorre essa polarização, os *recursos de natureza política fracassam*. Por discordarem quanto à matriz institucional a partir da qual a mudança política deverá ser atingida e avaliada, por não reconhecerem nenhuma estrutura supra-institucional competente para julgar diferenças revolucionárias, os partidos envolvidos em um conflito revolucionário devem recorrer finalmente às técnicas de persuasão em massa, que seguidamente incluem a força.

Com o decorrer da história da humanidade, as mudanças de paradigma na ciência mostraram ter muitas semelhanças com as revoluções políticas. Semelhantes aos estilos de governo, muitas vezes as posições políticas revelam a existência de modos de vida extremamente incompatíveis. Segundo essa comparação, os momentos de revolução não podem ser avaliados pelos processos comuns da ciência normal, porque eles dependem de um paradigma para a sua avaliação, o qual se encontra sendo questionado a respeito da sua validade enquanto um paradigma coletivamente aceito. Da mesma forma, isso pode ser indicado quando nos questionamos sobre a validade de documentos que apregoam a inocência de governos que passaram por períodos de barbárie.

A proposta de Kuhn trata tanto das grandes revoluções científicas quanto das pequenas. Kuhn (2011a) menciona que as características das revoluções científicas podem ser igualmente examinadas através do estudo de outros episódios que não foram tão obviamente revolucionários. As equações de Maxwell, por exemplo, afetaram um grupo bem mais reduzido do que as de Einstein, porém, não foram consideradas menos revolucionárias e, por esse motivo, encontraram resistência. Kuhn (2011a, p. 74) sugere "[...] a existência de revoluções grandes e pequenas, algumas afetando apenas os estudiosos de uma subdivisão de um campo de estudos. Para tais grupos, até mesmo a descoberta de um fenômeno novo e inesperado pode ser revolucionária".

Revolucionar, no sentido amplo, significa, em suma, trocar um paradigma por outro. A área de pesquisa passa de um período de dúvidas, incapacidades, insegurança, incertezas para um novo período de

desenvolvimento, confiança no paradigma, na sua capacidade de previsão, alcance explicativo e de sucesso na resolução de quebra-cabeças. A área de pesquisa se encontra em um novo período de ciência normal, assentando-se em momento de calmaria que, mais cedo ou mais tarde, em menor ou maior intensidade, voltará a ser de instabilidade. Assim gira a roda da estrutura das revoluções científicas.

Esta, em resumo, é a perspectiva de Kuhn da estrutura das revoluções científicas, do funcionamento da ciência, do progresso científico. Nos próximos dois capítulos apresentamos o surgimento da ciência cognitiva, um deles no contexto histórico e outro no contexto epistemológico, buscando averiguar o estatuto científico dessa nova área de pesquisa.

2

A ciência cognitiva em sua fase inicial: contexto histórico

Apresentação

Neste capítulo apresentamos o contexto histórico do surgimento da ciência cognitiva, o qual está dividido em quatro seções. Na primeira delas expomos as cinco principais características dessa área de pesquisa, estabelecidos em seu início: o uso de representações mentais; a crença de que os computadores eletrônicos são bons modelos explicativos de processos cognitivos; a decisão deliberada em não enfatizar certos fatores "demasiadamente complicados"; a interdisciplinaridade; e a consolidação de uma comunidade científica. Na segunda seção focalizamos os elementos históricos fundadores da ciência cognitiva, como os primeiros textos que fazem referência à área e à formação de seus primeiros membros. Na terceira seção tratamos de alguns fatos ocorridos nas conferências Macy e os temas que serviram de base para o surgimento oficial dessa área de pesquisa. Para fechar o capítulo, na última seção traçamos uma relação entre a Cibernética e o surgimento da ciência cognitiva.

2.1 Caracterização da ciência cognitiva em sua fase inicial

Na contemporaneidade, a ciência cognitiva obteve um papel de grande destaque ao abordar o tema da cognição atrelado à possibilidade de sua simulação a partir de aparatos eletrônicos.

Gardner (1996) sugere a presença de cinco características de maior importância, as quais podem ser consideradas sintomáticas da atividade concernente à ciência cognitiva. Quando todos ou boa parte estão presentes, pode-se assumir que se trata do tema da ciência cognitiva. Ainda que essas características possam ser consideradas sintomáticas, ressaltamos que elas fazem referência à fase inicial dessa área de pesquisa, ou seja, a década de 1940, conforme especificaremos mais adiante neste capítulo.

Em primeiro lugar, existe a crença, ao discutir a atividade cognitiva, de que é fundamental fazer o uso de representações mentais e criar níveis de análise capazes de separar o biológico do cultural.

Diversas foram as noções de representação mental propostas ao longo da história dessa área. Segundo Thagard (1998, p. 17): "A ciência cognitiva propõe que as pessoas têm procedimentos mentais que operam nas representações mentais para produzir pensamentos e ações." Tipos diferentes de representações mentais, como conceitos e regras, promovem e necessitam de diferentes tipos de procedimentos mentais. Assim, por exemplo, considerando pessoas familiarizadas com as representações da língua portuguesa e com os procedimentos para a formulação de sentenças, supõe-se que elas serão capazes de produzir orações com algum grau de satisfação. Já as pessoas familiarizadas com as representações da língua inglesa também serão capazes de formular orações, mas em outra língua, com suas próprias regras e estruturas. Neste exemplo, a ordem entre substantivo e adjetivo nas orações exige uma estrutura cognitiva distinta no uso das línguas.

Em segundo lugar, é fundamental, para essa área no momento de seu surgimento, a crença de que o uso de sistemas eletrônicos, autômatos como os computadores digitais, por exemplo, é indispensável na realização de estudos cognitivos. Além disso, entende-se que tais sistemas poderiam ser a melhor forma para criar modelos ética e funcionalmente viáveis para a explicação de processos cognitivos, uma vez que são capazes de replicar ou de simular tais processos.

Uma das melhores formas de desenvolver estruturas teóricas é formando e testando modelos computacionais que pretendem ser semelhantes às operações mentais. Em sua forma ideal e forte, na ciência cognitiva, os modelos computacionais e a experimentação psicológica

andam lado a lado, possuindo uma equivalência entre o que é representado digitalmente, através das máquinas, e o que acontece no processo cognitivo.

Diante das dificuldades materiais e éticas para a manipulação da mente, cujo substrato físico é o cérebro, como defendem a quase totalidade dos cientistas cognitivos, pesquisadores como Herbert Simon (1969) adotam a hipótese de que certos processos cognitivos podem ser simulados ou explicados a partir de modelos computacionais. Tais sistemas são *hardwares* que processam informação, ou seja, que executam *softwares,* assim como os processos cognitivos são processamentos de informação instanciados, geralmente, em alguma base física.

De acordo com Dupuy (1996), modelos tratam de uma idealidade formalizada e matematizada, a qual tem a função de sintetizar um sistema de relações entre elementos que podem ser substituídos por outros elementos análogos ou diferentes, sem que o modelo seja alterado. Essa forma de abordagem norteou grande parte dos programas de pesquisa em Inteligência Artificial. O uso dos computadores e as definições de Inteligência Artificial e de ciência cognitiva estão intimamente relacionadas, conforme Boden (1990, p. 1, tradução nossa):

> Inteligência Artificial (IA) é algumas vezes definida como o estudo de como construir e/ou programar computadores para lhes permitir fazer o tipo de coisas que mentes podem fazer [...] Porém, muitos outros preferem uma definição mais controversa, vendo a IA como a ciência da inteligência em geral – ou, mais precisamente, como o núcleo intelectual da ciência cognitiva. Como tal, seu objetivo é prover uma teoria sistemática que explique (e talvez nos habilite replicar) não só as categorias gerais da intencionalidade e a diversa psicologia das criaturas terrestres, mas também o conjunto inteiro das possíveis mentes.

Assim, os estudos da ciência cognitiva estão intrinsecamente relacionados ao desenvolvimento da Inteligência Artificial, envolvendo o processamento de informação, a manipulação de símbolos e a noção de representação. Interpretação semelhante também pode ser verificada a partir da definição de ciência cognitiva oferecida por Dawson (2002, p. 13, tradução nossa):

> Ciência cognitiva é um estudo intensamente interdisciplinar de cognição, percepção e ação. Ela está baseada na hipótese de que cognição é processamento de informação, em que processamento de informação é geralmente interpretado como manipulação baseada em regras de estruturas de dados que estão armazenadas numa memória. Como resultado dessa hipótese, um objetivo básico da ciência cognitiva é identificar a arquitetura funcional da cognição – o conjunto primitivo de regras e representações que medeiam o pensamento.

O terceiro aspecto, já apontado acima, surge da crença de que há muito a se ganhar com estudos e trabalhos interdisciplinares. Por um lado, a complexidade do tema exige a participação de diversas áreas de pesquisa. Por outro lado, a maioria dos cientistas cognitivos na fase inicial desta disciplina era advinda de áreas de pesquisa distintas, como física, neurociências, antropologia, psicologia, matemática. Imaginava-se, então, que haveria um estreitamento entre os limites das disciplinas até a consolidação de uma só ciência cognitiva unificada, com um paradigma próprio.

Existem diversas perspectivas e métodos que os pesquisadores de diferentes áreas trazem para os estudos da ciência cognitiva. Os neurocientistas, por exemplo, desempenham experimentos controlados e suas observações são direcionadas, habitualmente, a processos e relações cerebrais. Com o desenvolvimento das tecnologias recentes, tornou-se possível o uso de aparelhos para observar o que acontece nas mais diversas regiões cerebrais, enquanto se está executando várias tarefas cognitivas. A antropologia cognitiva, por sua vez, expande o exame do pensamento humano para considerar o impacto dos diferentes ambientes culturais na cognição. Os antropólogos investigam, por exemplo, as semelhanças e diferenças entre as culturas, nas palavras que representam cores.

No tocante aos filósofos, geralmente eles não fazem observações empíricas ou trabalham com modelos computacionais. Assim, caberia à filosofia, dentre outras coisas, lidar com tópicos fundamentais subjacentes às abordagens experimentais e computacionais da cognição e refletir sobre as bases epistemológicas dos conceitos arrolados nas pesquisas ou das consequências éticas das pesquisas e descobertas científicas. Nesse sentido,

a filosofia seria muito mais uma espécie de analista crítica das abordagens teóricas propostas na ciência cognitiva. Entretanto, há divergências no tocante ao seu papel. Para muitos pensadores, a filosofia também deveria contribuir para propostas de abordagens explicativas a respeito da cognição. Temas abstratos, como a natureza das representações, as relações entre a mente e corpo e os critérios de cientificidade da ciência cognitiva, dentre outros, são temas importantes para o filósofo que se preocupa com a ciência cognitiva ou com os estudos dos processos cognitivos. Embora não tenhamos exemplificado todas as áreas específicas, destaca-se a importância que a convergência teórico-experimental dessas áreas tem sobre os avanços científicos em relação à cognição, conforme pode ser conferido na obra de Thagard (1998), por exemplo.

O quarto aspecto da ciência, em sua fase inicial, menciona a decisão deliberada por grande parte dos cientistas em não enfatizar alguns fatores que possuem certo grau de relevância, cuja inserção, todavia, complicaria demasiadamente o trabalho dos cientistas cognitivos. Estão incluídos, entre esses fatores, a influência das emoções, o contexto histórico e cultural, além do pano de fundo em que ocorrem as ações ou a cognição.

O quinto aspecto afirma que outro ingrediente fundamental, semelhante à proposta de Kuhn (2011a), é a consolidação de uma comunidade científica. Neste bojo também é importante a construção de uma agenda de questões e conjuntos de preocupações, os quais são capazes de nortear as pesquisas científicas e que há muito tempo inquietam os pensadores ocidentais que se debruçaram sobre o problema da cognição.

Estes cinco aspectos servem de fundamento para a consolidação da existência de um possível paradigma da ciência cognitiva em sua fase inicial. Assim, são os três primeiros aspectos princípios metodológicos do paradigma que asseveram o uso de representações mentais, a crença de que é fundamental o uso de computadores eletrônicos e a decisão deliberada em não enfatizar certos fatores. Os outros dois últimos aspectos indicam a formação desejada aos cientistas cognitivos e o papel fundamental que a comunidade científica exerce no período da inauguração e vida de uma área de pesquisa. Em outras palavras, em termos teóricos, houve a tentativa de estabelecer pressupostos metodológicos, filosóficos, até metafísicos, um arcabouço conceitual, regras de conduta nas pesquisas, teorias científicas sobre um dado conjunto de fenômenos no mundo, além do estabelecimento

de uma comunidade científica coesa, articulada e unificada em torno dessa matriz disciplinar. No entanto, será que isso foi possível na prática? É esta questão que propomos investigar no restante desta obra.

2.2 Contexto histórico para o surgimento da ciência cognitiva

A ciência cognitiva nasce com o intuito de responder empiricamente, através do uso do computador, questões que os primeiros filósofos já haviam formulado sobre o estudo da mente. Dar uma definição precisa desse novo campo de pesquisa em sua fase inicial não é uma tarefa simples. Trata-se de uma área de pesquisa recentemente constituída e, talvez por isso, seu campo de atuação, método e objeto ainda não estavam bem delimitados naquele momento. Gardner (1996, p. 19-20) define a ciência cognitiva como

> [...] um esforço contemporâneo, com fundamentação empírica, para responder questões epistemológicas de longa data – principalmente aquelas relativas à natureza do conhecimento, seus componentes, suas origens, seu desenvolvimento e emprego. Embora o termo ciência cognitiva seja às vezes ampliado, passando a incluir todas as formas de conhecimento – tanto animado como inanimado, tanto humano com não humano – aplico o termo sobretudo a esforços para explicar o conhecimento humano. Interessa-me saber se questões que intrigavam nossos ancestrais filosóficos podem ser definitivamente respondidas, ilustrativamente reformuladas, ou permanentemente abandonadas. Hoje a ciência cognitiva tem a chave para decidir.

Segundo Churchland (1984, p. i, tradução nossa), "[...] uma das funções principais da ciência cognitiva é estudar quais são os elementos básicos da atividade cognitiva e como eles podem ser implementados em sistemas físicos reais". Desse modo, a preocupação inicial da ciência cognitiva não é a caracterização de estados e faculdades mentais em geral, mas sim a caracterização e a simulação de processos cognitivos.

Existe certa variedade de datas importantes sugeridas por historiadores para o surgimento da ciência cognitiva. Dupuy (1996) destaca os anos entre 1946 e 1953, quando ocorreram dez conferências que

reuniram inúmeros pensadores da época, sendo as nove primeiras realizadas no Hotel Beekman, 575 Park Avenue, em Nova York, e a última, no Hotel Nassau de Princeton, Nova Jersey. Essas conferências entraram para a história com o nome de Conferências Macy, graças à organização efetuada pela fundação filantrópica Josiah Macy Jr. Tais reuniões eram compostas por matemáticos, lógicos, engenheiros, fisiologistas e neurofisiologistas, psicólogos, antropólogos e economistas. Eles tinham como objetivo fundamentar uma ciência geral capaz de tratar de aspectos cognitivos.

Em princípio, esse grupo adotou o nome de cibernética. Com o passar dos anos, novos nomes foram utilizados, até estabelecer-se com a expressão ciência cognitiva para nomear a área de pesquisa. A palavra cibernética vem do grego e significa a arte de governar. Não é à toa que Wiener, um dos seus fundadores, na década de 1940, escolhe este nome para uma das mais novas e promissoras áreas de pesquisa naquele momento. Na concepção de Wiener (1961), a cibernética consiste no domínio todo da teoria da comunicação e do controle, seja na máquina ou no animal. Ruyer (1992) afirma que a cibernética é a ciência do controle, por meio de máquinas de informação, sejam elas naturais, como as máquinas orgânicas, sejam elas artificiais. Os cibernéticos, ao estudar processos cognitivos, acreditavam que seu trabalho devia ser interdisciplinar, dada a complexidade do objeto de estudo, por meio da utilização de modelos computacionais.

Os eventos históricos a que fazemos menção estão relacionados a um movimento histórico e intelectual de sua época. Embora as primeiras conferências tenham tido início apenas em 1946, algumas outras datas e eventos são importantes, para que possamos compreender as ideias daquele momento. Em 1942, ou seja, em plena Guerra Mundial, foi promovida em Nova York uma conferência, sob a tutela da anteriormente citada fundação Josiah Macy Jr., cujo objetivo era discutir os problemas atrelados à inibição no sistema nervoso central. Essa reunião fez com que grandes pensadores da época entrassem em contato, possibilitando os futuros encontros que dariam origem às Conferências Macy. Segundo Dupuy (1996), esse evento de 1942 contou com a presença dos fisiologistas Arturo Rosenblueth e Warren McCulloch, com o casal de antropólogos Gregory Bateson e Margaret Mead, além do diretor médico da fundação, Frank Fremont-Smith. McCulloch, ao notar certa relação entre as ideias sugeridas no

evento e aquelas que ele desenvolvia com Pitts, propôs a Fremont-Smith que organizasse uma série de conferências, no modelo das que a fundação promovia, para tratar de temas médico-sociais muito variados, as quais deveriam ser realizadas apenas depois do fim da Segunda Guerra Mundial.

Conforme destaca Dupuy (1996), após esse encontro, aconteceram inúmeros intercâmbios entre o grupo de McCulloch, em Chicago, e o de Wiener, no Massachusetts Institute of Technology (MIT). Em 1943, eram publicados, independentemente, dois dos artigos que acabariam compondo a base dos pressupostos do que mais tarde seria conhecido como movimento cibernético. O primeiro foi assinado por Arturo Rosenblueth, Norbert Wiener e Julian Bigelow, intitulado “Behavior, Purpose and Teleology”. O segundo, intitulado “A logical calculus of the ideas immanent in nervous activity”, foi escrito pelo neuropsiquiatra McCulloch em parceria com o matemático Walter Pitts.

Wiener e Bigelow trabalhavam, ao longo da guerra, nos problemas teóricos levantados pela defesa antiaérea. O problema central da defesa antiaérea é que, como o alvo é móvel, é preciso prever a sua posição final com base em uma informação parcial da trajetória. Os problemas da defesa antiaérea ajudaram a fundamentar um dos ingredientes básicos da futura cibernética, o conceito de retroalimentação (*feedback*), inerente às noções entre a ação efetiva (*output*), a entrada de informação (*input*) e os resultados projetados. Graças a essa noção, os objetos seriam capazes de mudar as suas relações, conforme ocorresse a presença de novos estímulos e respostas.

Dupuy (1996) sugere que, com o texto fundador de Rosenblueth, Wiener e Bigelow, tem-se uma concepção metodológica que recomenda que se deixe de lado a natureza física dos constituintes e de suas relações, a fim de abstrair a sua forma. Por exemplo, um empreendimento fundamental da cibernética trataria de descobrir um mesmo dispositivo formal de *feedback* de um animal e de uma máquina, embora o primeiro seja materialmente composto por proteínas, e o segundo, por sistemas eletrônicos. Esse procedimento pretendia propor uma teoria unificada entre os seres vivos e as máquinas.

O segundo artigo fundador também foi publicado em 1943, pelo neuropsiquiatra McCulloch e pelo matemático Walter Pitts. Esses dois pesquisadores preconizavam investigar os mecanismos materiais e lógicos que compõem a atividade cognitiva, baseada no processamento

cerebral. O título por eles escolhido foi republicado em 1965, em uma coletânea com seus principais artigos, com o nome de *Embodiments of Mind* (McCULLOCH, 1965). A preocupação de McCulloch era, por assim dizer, de cunho filosófico: buscava descobrir como sabemos e como desejamos. Para ele, os eventos mentais não são compreendidos como submissos a uma entidade controladora abstrata. Ao contrário, são explicados a partir de conjunções sinápticas. Com esse objetivo e tendo em vista a ideia da encarnação da mente (tentativa de identificação entre mente e cérebro), McCulloch e Pitts (1943) constroem uma rede neural artificial, ainda seguidora de regras lógicas, tomada como um modelo do cérebro. Ressalta Dupuy (1996, p. 53):

> McCulloch, na verdade, introduz o "estudo comportamental dos fenômenos naturais", caro a Wiener, Rosenblueth e Bigelow, no interior do cérebro. Sem dúvida, o "conteúdo" daquilo mesmo que é capaz de comportamento é agora considerado pertencente à esfera de um procedimento científico, mas esse conteúdo se descreve a si próprio em termos de comportamento de unidades menores, no "interior" das quais não se pode pensar em penetrar e que só são consideradas em suas relações com seu ambiente, ou seja, como operadores que transformam *inputs* em *outputs*: os neurônios.

A posição de McCulloch reflete alguns dos pensamentos desse primeiro momento da cibernética, que pode ser resumido através da equivalência lógica entre os artefatos e as máquinas artificiais com todo e qualquer ser lógico-matemático instanciado na matéria de algum organismo. Segundo McCulloch e Pitts (1943), grande parte das atividades cognitivas poderiam ser descritas em termos de conexões, e estas, em termos de proposições lógicas. Logo, por transitividade, tais atividades poderiam ser descritas por meio de proposições lógicas. Nas palavras de McCulloch (1955, p. 38, tradução nossa): "Quanto mais aprendemos sobre os organismos, mais somos levados a concluir que eles não são simplesmente análogos às máquinas, mas são máquinas."

Nesse sentido, baseados na perspectiva epistemológica de Kuhn (2011a), poderíamos dizer, em um primeiro olhar que, embora a comunidade científica da ciência cognitiva ainda estivesse sendo consolidada, além dos centros de pesquisa, havia aspectos metodológicos

que buscavam agrupar e nortear os seus primeiros partidários. Dentre os inúmeros aspectos, Wiener, Rosenblueth e Bigelow sugeriam um enfoque metodológico, enquanto Rumelhart e McClelland apresentavam uma metáfora sobre o funcionamento da mente, que se unia às noções gerais sobre as máquinas digitais da época.

Ainda durante 1943, a aproximação dos grupos fundadores concretizou-se pela vinda de Pitts ao MIT, junto a Wiener. Pitts e Wiener deveriam, a partir desse momento, percorrer com frequência o triângulo entre Cambridge, Cidade do México e Chicago.

Paralelamente, Wiener mantinha uma forte relação de intercâmbios com Von Neumann, para tratar de questões que envolviam as analogias entre organismos e máquinas. Eles chegaram à conclusão de que as coisas já estavam maduras para uma institucionalização das pesquisas nessa área, sob a forma de revista, de uma sociedade científica e até de um centro de pesquisa. Nessa ocasião, seguindo as elucidações de Kuhn (2011a) para o estabelecimento de um paradigma científico, desenvolviam-se os primeiros exemplares e periódicos destinados à nova comunidade científica.

Conforme Dupuy (1996), em 1945, no *Institute of Advanced Studies de Princeton*, onde Von Neumann trabalhava, foi organizado um encontro que contou com a participação de Von Neumann, Wiener, Pitts, Goldstine, McCulloch e Lorente de No, para tratar da concepção do Eniac (em português, computador integrado numérico eletrônico), o primeiro computador digital eletrônico de grande escala. Após sair desse encontro, Wiener estava convencido de que a engenharia e a neurologia estavam intimamente relacionadas e de que havia uma necessidade de se organizar um programa permanente de pesquisa, para tratar desse assunto.

Entretanto, devido à forma como o instituto de Von Neumann estava reticente a deixá-lo sair da sua instituição, permitiu que ele construísse ali o seu computador ultrarrápido, impossibilitando que Von Neumann se dirigisse ao MIT com Wiener para a construção do tão sonhado centro. Assim, a primeira cibernética nunca conseguiu dispor de um centro de pesquisa apenas seu, encontrando fortes empecilhos durante a consolidação de seu paradigma.

As Conferências Macy deveriam, de algum modo, suprir essa falta de um centro de pesquisa. Terminada a guerra, a Fundação Macy retomou a sugestão feita por McCulloch, em 1942, e o encarregou de montar uma série de conferências sobre essas ideias nascentes. Tais encontros poderiam ser oportunidades essenciais para o fortalecimento da comunidade científica em torno de uma matriz disciplinar, angariando novos adeptos, estabelecendo um corpo de definições elementares do paradigma, de um conjunto de teorias testáveis empiricamente, dentre outros fatores, a fim de propiciar o direcionamento das pesquisas e o desenvolvimento da área através de novas descobertas a abordagens explicativas dos fenômenos investigados, quais sejam, os processos cognitivos.

2.3 As Conferências Macy

O primeiro encontro organizado pela Fundação Macy aconteceu em março de 1946, em Nova York, intitulada *Feedback Mechanisms and Circular Causal Systems in Biological and Social Systems*. Conforme Dupuy (1996, p. 85):

> Segundo os princípios da fundação, tratava-se de reunir em intervalos regulares (em geral, a cada seis meses) um pequeno grupo de cerca de vinte pesquisadores, membros oficiais do ciclo em questão, os quais podiam acrescentar até cinco "convidados". A ênfase era dada muito mais aos intercâmbios e às discussões do que às exposições formais.

O título da primeira conferência fazia referência aos trabalhos de Wiener e de McCulloch, dada a importância atribuída aos neurônios na sua abordagem. Houve também a inclusão dos "sistemas sociais", resultante da iniciativa de Bateson. A maior parte dos participantes era composta por psicólogos, sociólogos, matemáticos e antropólogos.

A segunda Conferência Macy, com o novo título de *Teleological Mechanisms and Circular Causal Systems*, aconteceu em outubro do mesmo ano da primeira. Imediatamente após essa conferência, foi realizado, pelo mesmo grupo, um simpósio para a Academia de ciências de Nova York, sobre o tema *Teleological Mechanisms*. As ideias do grupo foram apresentadas nessas duas ocasiões, afirmando-se que seu objetivo era de consolidar um

quadro de conceitos para as pesquisas científicas nas ciências da vida. Segundo Dupuy (1996), foi nesse encontro que Wiener expôs a um grupo de cientistas, pela primeira vez, as noções de mensagem, comunicação, informação, *feedback* e autômato.

Em 1947 foi realizada a Terceira Conferência Macy, com o mesmo título da segunda. Foi nessa ocasião que Wiener sugeriu o termo "cibernética", para consolidar uma unidade ao grupo e às suas ideias. No ano seguinte, 1948, Wiener publicou o seu famoso livro de mesmo nome, *Cybernetics: or the control and communication in the animal and the machine* (WIENER, 1948). No mesmo ano de sua publicação, são promovidas as duas últimas conferências do primeiro ciclo, ambas intituladas "Circular Causal and Feedback Mechanisms in Biological and Social Systems".

O primeiro ciclo de conferências Macy, aos olhos da proposta de Kuhn (2011a), não se consolida enquanto fazendo parte de um estado de ciência normal. Embora a interdisciplinaridade seja considerada um elemento fundamental dessa área de pesquisa, acaba por dificultar a criação de exemplares e princípios metodológicos bem esclarecidos. A cibernética, nesse primeiro momento, ainda buscava atrair um número de partidários. Por esse motivo, ainda que carecesse de paradigmas propriamente ditos já bem estabelecidos, a ciência cognitiva, em seu gérmen, deve ser caracterizada em momento de pré-ciência. O primeiro ciclo de conferências ainda não conseguiu estabelecer um paradigma para a área em construção, sequer, portanto, um paradigma dominante e independente, com um conjunto de conceitos bem elaborados, pressupostos metodológicos e regras de pesquisa claras. Talvez um pouco mais de amadurecimento, em um novo ciclo de conferências, poderia propiciar tal criação, tornando essa área de pesquisa um tanto mais inclinada a um momento de ciência normal, possibilitando a criação de uma comunidade coesa em torno de uma matriz disciplinar, propondo teorias explicativas dos processos cognitivos.

O segundo ciclo de Conferências Macy ocorreu entre 1949 e 1953. Estas foram as últimas cinco conferências sobre o tema da cibernética. Dupuy (1996) destaca que as atas dessas últimas conferências são marcadas, em 1951, pela ausência de Wiener e de Von Neumann, e, no ano seguinte, pelo desaparecimento de seus nomes da lista de membros do ciclo. A história desse período encerra-se com o rompimento de todas as relações de Wiener com o grupo.

A sexta Conferência Macy (24 e 25 de março de 1949) iniciou-se com uma discussão conjunta, cujo objetivo era investigar os mecanismos cerebrais responsáveis pela memória, especialmente a humana. O respectivo problema havia sido introduzido a partir das considerações anteriores de Von Neumann, conforme consta em Macy 6 (1950, p. 12, tradução nossa): "[...] os atuais 10^{10} neurônios [estimados num cérebro], usados como relés simples, são totalmente insuficientes para responder pelas habilidades humanas". Nesse momento da história da cibernética, entendeu-se que os pesquisadores deveriam procurar por estruturas menores no interior dos neurônios, as quais funcionassem de maneira "digital". Observamos a seguinte citação relatada pelo editor das conferências, Heinz von Förster, em Macy 6 (1950, p. 12, tradução nossa), de uma fala de McCulloch:

> *McCulloch*: [...] von Neumann é muito insistente de que, sejam quais forem os itens subjacentes à propriedade dos neurônios de serem dispositivos de tipo tudo ou nada, eles ainda devem ser quantizados ou digitais ou lógicos em sua estrutura. Você simplesmente não pode, numa coisa com a dimensão de nossos cérebros, se safar com dispositivos analógicos de nenhum tipo. É simplesmente impossível manejar informação suficiente dessa forma analógica.

Nesse momento, era colocado em discussão o que deveria ser procurado nas dimensões dos neurônios e qual o seu grau de complexidade (MACY 6, 1950). Os pensadores da época chegaram a algumas possíveis respostas. Inicialmente, McCulloch excluiu dessas possíveis respostas um modelo químico-mecânico elaborado por Ward e Katz. De acordo com McCulloch (MACY 6, 1950), esses dois cientistas haviam proposto um modelo para a memória no qual ela seria fruto de alterações temporárias na estrutura proteica da membrana das células neuronais, porque seria possível pensar o disparo dos neurônios como um tipo de reação físico-química e não apenas elétrica, como os impulsos nervosos que utilizam a mesma fonte de energia que as contrações musculares. Os disparos dos neurônios poderiam ser concebidos, pois, como uma alteração mecânica similar à dos músculos, literalmente, um ato de esticar da região da sinapse até o ponto de tocar um neurônio vizinho. Nesse processo, haveria uma alteração na estrutura química das células neuronais em questão, o qual,

segundo os relatos de Heinz von Förster em Macy 6 (1950, p. 13, tradução nossa), formaria:

> [...] uma espécie de rede na membrana [das células neuronais, e] esta rede, no momento do impulso nervoso, seria temporariamente alterada [...] poderia conduzir a especificidades na resposta de uma membrana proteica em relação a uma membrana proteica adjacente na sinapse.

Entretanto, esse modelo é considerado, para McCulloch, em Macy 6 (1950), insuficiente, uma vez que ele conduz a uma soldagem de um neurônio ao próximo, causando uma continuidade que contradiz a ideia, desse período, da tentativa de digitalização dos processos cerebrais.

Wiener também argumenta que a memória poderia ser resultante das alterações químicas nos neurônios, mas de maneira diferente das especulações de Ward e Katz. Para Wiener e os reducionistas, há uma ideia de que a memória estaria corporificada no nível celular, relata Heinz von Förster em Macy 6 (1950, p. 16, tradução nossa): "O que precisamos é de um processo muito mais complexo, em que os elementos são muito mais complexos para o que acontece na sinapse e no corpo celular". Essa complexidade não poderia ser alcançada a partir de alterações limitadas a um local específico, mas apenas através da alteração em conjunto de um grande número de ligações sinápticas. Wiener, ao longo da conferência Macy 6 (1950), partiu da hipótese de que a memória seria fruto de uma alteração no limiar de ativação de um grupo de sinapses. Uma vez que certos processos químicos poderiam influenciar a quantidade de energia recebida necessária para o disparo de um neurônio ou de um grupo deles, seria possível imaginar que a memória poderia ser pensada com base em sua incorporação em uma rede neural, "separada" de outras redes cerebrais por diferenciações, quimicamente controladas, do limiar de ativação daquela rede, garantindo a integridade da memória e explicando possivelmente os mecanismos que permitem o acesso a ela. A variabilidade dos limiares de ativação poderia explicar os processos de aprendizagem, na medida em que as suas mudanças poderiam representar uma alteração na quantidade ou nos "dados" armazenados naquela rede neural. Wiener, em Macy 6 (1950, p. 17, tradução nossa), acrescenta:

> O mecanismo de limiar variável me parece estar em forte conexão com o problema da memória. Eu acredito que possuímos todas as indicações de que várias coisas no corpo, químicas ou não, alteram-no: que, em outras palavras, temos uma evidência da existência da memória, pela existência de aprendizado, isto é, de uma fina estrutura celular de alterabilidade de limiares de acordo com padrões específicos [...] não precisa ser uma variabilidade fixa no espaço, mas pode com efeito ser uma variabilidade no tempo.

Pitts, partindo de um pensamento lógico-matemático, durante a sexta conferência Macy (MACY 6, 1950), examinou minuciosamente as respostas possíveis para o problema dos neurônios propostos por seus colegas. Segundo Pitts, em Macy 6 (1950, p. 22, tradução nossa), pode-se distinguir “[...] três estágios da complicação da relação entre a entrada e a saída [de informação] da célula”. A primeira possibilidade, a mais simples, é aquela na qual as sinapses agem juntas. Esse primeiro cenário é considerado pelo seu autor bem pouco realista. A segunda possibilidade sináptica é aquela na qual o requisito para o disparo das células é que um grupo de sinapses próximas disparem em conjunto. Nesse cenário, em vez de se ter em vista uma célula, considera-se um grupo de células como agente do disparo. A terceira possibilidade para a estruturação das sinapses parte da ideia de combinações não aditivas; tal tipo de combinação conduz a uma estrutura sináptica muito rígida e específica: “[...] uma célula na qual alguma combinação arbitrária fixa de sinapses aferentes – e somente aquela combinação – irá disparar” (MACY 6, 1950, p. 22, tradução nossa). Pitts classifica as possibilidades em graus de eficiência no processamento de informações. Assim, a segunda possibilidade de organização é tomada como mais eficiente que a primeira, pois nela há a possibilidade de discriminação maior, uma vez que cada grupo de células determina de antemão certos padrões de organização dos impulsos nervosos, enquanto exclui outros. O terceiro caso representa uma maior eficiência, pois consiste em uma única forma de organização. Pitts (MACY 6, 1950) chega a especular sobre uma quarta possibilidade de organização sináptica, levando em conta a organização temporal, na qual os impulsos poderiam ser resultantes de tempos diferentes e combinações diferentes.

O fisiologista Gerald, por sua vez, durante a sexta conferência Macy, chamará a atenção para o problema da estrutura das redes neurais.

Segundo Gerald, em Macy 6 (1950, p. 18, tradução nossa), há problemas em se atribuir alguma fixidez às organizações dos neurônios:

> [...] Uma coisa arrebatadoramente impressionante neste trabalho é a completa fluidez – e eu estou falando de fluidez estrutural – de um neurônio adulto sob condições normais e sua extensa disrupção e reconstrução sob leves condições patológicas. Nesse momento, em cada um de nossos cérebros um neurônio não está lá sentado como uma figura na lousa, tal como ordinariamente pensamos nele. Cada um está projetando pseudópodes, retraindo suas fibras, movendo-se para frente e para trás, esticando e encolhendo e movendo-se de um lado para o outro. Todas as vezes que se assiste a filmes dessas coisas, não importa quão frequentemente, fica-se impressionado com o fato de que dificilmente há algo mais do que um fino gel.

Essa primeira discussão terminou de maneira inconclusiva, mas tratou de estimular o ambiente para as discussões que ainda estavam por vir ao longo do encontro. Esse debate inicial levantou uma problemática referente ao processamento de informação realizado pelo cérebro e seu armazenamento. Toda a controvérsia girou em torno de qual seria a unidade de processamento da informação, se são neurônios inteiros, parte de seus corpos, se são diversos neurônios, se são sinapses ou um grupo de sinapses, além da sua estrutura, se são fixas ou mutáveis, aditivas ou não aditivas. Nessa discussão, a noção de troca de informação serve como um elemento fundamental para a modelização do funcionamento neuronal e da memória.

Ainda durante a sexta conferência Macy, foi apresentado pelo psicólogo e pesquisador do Laboratório de Eletrônica Naval de San Diego, John Stroud, um modelo da percepção humana. Essa conferência foi intitulada "The Psychological Moment in Perception". Stroud começou sua argumentação comentando o trabalho de Kenneth Craik. Este pesquisador britânico buscou pensar, em termos cibernéticos (com base na noção de *feedback*), o funcionamento de uma arma antiaérea. Sua preocupação, entretanto, não era a mesma de Wiener (1948), que procurava, nesse tipo de dispositivo, um mecanismo aplicável, por analogia, ao ser humano. Em vez disso, Craik se perguntava sobre como pensar a ação do operador humano de tal arma. Em teoria, uma arma antiaérea é um dispositivo que,

ao receber informações sobre seu alvo, como velocidade e direção, as utiliza para prever sua posição, num momento futuro, e mira no ponto que, calculado a partir das informações recebidas, dispara o projétil a fim de encontrar o alvo. Todavia, como nem sempre esse cálculo é preciso, porque a arma pode errar o tiro, as informações relativas a esse erro realimentam o próprio artefato, que passa a calcular sua margem de erro, a quantidade pela qual não alcançou o alvo e a margem com que deve compensar os disparos futuros. O *feedback* se dá, portanto, em função do erro. A operação da arma deve ser corrigida automaticamente, a partir da recepção dos erros cometidos.

O instrumento concebido nesse experimento é reflexo da tecnologia da época, ou seja, esse tipo de arma até então não era completamente automático, sendo necessário que um operador humano decidisse quando e contra o que atirar:

> [...] temos o operador humano cercado por todos os lados por mecanismos conhecidos muito precisamente, e a questão que surge é "que tipo de máquina colocamos no meio?" Craik chegou à conclusão de que o operador humano é um servomecanismo intermitente (MACY 6, 1950, p. 28, tradução nossa).

Isso pode significar duas coisas: em primeiro lugar, ao contrário de um servomecanismo tradicional, que funciona sozinho recebendo sua informação continuamente e dando sua informação de entrada e resposta de saída continuamente, o operador humano funciona de forma intermitente, isto é, ele toma suas decisões somente após um certo intervalo de tempo, algo em torno de dois a três milissegundos, durante o qual permanece inativo ou em estado de processamento de informação. Em segundo lugar, em vez de basear-se apenas no erro, isto é, de orientar sua ação sempre em razão de quanto falta para atingir sua meta, o operador humano o faz com base na predição da movimentação do alvo (MACY 6, 1950).

Curioso em decifrar as razões das observações de Craik, Stroud, em Macy 6 (1950), montou o seguinte experimento: tem-se um instrumento, um painel com um ponteiro e uma linha vertical, um botão que gira e controla o deslocamento horizontal do ponteiro. É tarefa de um operador humano corrigir o ponteiro, que se desloca por conta própria, de sorte que ele se mantenha sobre uma linha vertical. Notemos que esse

tipo de controle exercido sobre o ponteiro, através do botão, pode ser regulado para controlar o deslocamento do ponteiro, ou sua velocidade, ou sua aceleração, ou ainda uma combinação dos três parâmetros. Quer dizer, a quantidade de deslocamento efetuada sobre o botão corresponderá diretamente ao deslocamento sofrido pelo ponteiro, ou a uma mudança em sua velocidade de deslocamento, ou à alteração da taxa de aceleração do movimento do ponteiro.

Stroud (MACY 6, 1950) chegou a alguns resultados, a partir deste experimento: o sucesso em manter o ponteiro sobre a linha vertical, mesmo em casos demasiadamente complicados, está relacionado com uma diferença na capacidade de *feedback* humana em relação à dos servomecanismos tradicionais, que, por possuírem "um curso de ação fixo", são incapazes de passar no teste desse experimento. Mas, se o *feedback* realizado pelo operador humano que passa por tal experimento não funciona tal qual o *feedback* dos servomecanismos tradicionais, como ele funcionaria? Destaca Stroud em Macy 6 (1950, p. 32, tradução nossa):

> [...] quando se analisam os registros [do experimento, vê-se que o controlador humano] faz seus ajustes primeiro em termos de quanto ele errou na simples escala de deslocamento. Um pouco depois, começa a ser aparente que o operador está agora fazendo seus ajustes parcialmente num sistema simples [deslocamento] e parcialmente sobre a base de quanto a velocidade está envolvida. Um pouco depois ele pode vir a incluir a aceleração. Se então você introduzir alguma mudança brusca, você verá que a solução dele vem abaixo, ele [então] volta para o seu conjunto de ajustes original e passa através das várias ordens de derivativos [deslocamento, velocidade e aceleração], produzindo soluções sucessivamente melhores para os problemas [...] Ele faz isso de uma forma muito peculiar. Ele não o faz continuamente. Tipicamente, ele [o operador humano] é um corretor de meio-ciclo, pois ele faz correções a cada meio segundo ou a cada terço de segundo. Suas correções estão predeterminadas.

O controlador humano, ao contrário dos servomecanismos, que agem continuamente a partir do erro de sua ação anterior, segue um padrão predeterminado, obtido pela observação dos resultados de sua ação, mas a executando em intervalos de tempo intermitentes. Nesse

caso, o padrão consiste em aumentar o grau de controle sobre o ponteiro, incorporando sucessivamente o deslocamento, a velocidade e a aceleração, como parâmetros da ação de controle. A eficácia dos resultados desse padrão de controle é checada a cada um terço ou metade de segundo; caso haja sucesso, o controlador persiste em seu padrão de ação; caso haja fracasso, ele inicia todo o processo novamente, desde o início. De acordo com Stroud em Macy 6 (1950, p. 33, tradução nossa), "[...] a informação que determina sua próxima ação corretiva não é obtida durante o curso da própria ação corretiva". Portanto, o *feedback* próprio à ação humana orientada é de tipo preditivo, segundo Stroud em Macy 6 (1950, p. 34, tradução nossa): "O que ele [o homem] fará, foi decidido por predição [...] quando ela [a ação] se inicia, não há nada que ele possa fazer [para corrigi-la]. Se, enquanto isto, a informação muda, ele nada pode fazer".

Em vez de checar continuamente o resultado de sua ação, o ser humano projeta um padrão de ação até então bem-sucedido, alterando-o apenas em caso de erro, que, quando percebido, não serve como informação corretiva da ação em vias de execução, mas somente da próxima ação. Justamente por isso, o humano age de forma preditiva e, não possuindo a instantaneidade das máquinas, sua ação baseia-se na repetição de padrões bem-sucedidos. Essa condição só é possível porque o ser humano dispõe de uma memória. Ele não checa a todo instante o resultado de sua ação, porém, antes, "prevê" se um determinado modo de ação será eficaz, quer dizer, repete um padrão que ele sabe, por experiência própria, que pode conduzir a determinados resultados.

A partir desse experimento, que corrobora, em certa medida, a proposta de Craik, Stroud tenta derivar uma teoria quântica ou digital da percepção. Para Stroud, a razão para o caráter projetivo da ação humana orientada por fins reside na incapacidade de processar informações de maneira contínua:

> Suponha que tenhamos uma tarefa em que nosso operador recebe sua informação por meio de seus olhos; números muito grandes de fótons são absorvidos na retina a taxas estatisticamente estáveis, de forma que podemos falar deles como sendo recebidos continuamente. Eis aonde a informação entra no organismo humano, e ela entra, para todos os propósitos, continuamente. Quando analisamos o que sai do organismo, todo conjunto

> de observações de suficiente sensibilidade que foram até agora analisadas mostraram periodicidades de baixa frequência, frequências da ordem de duas ou três [respostas] por segundo. Há um período da ordem de um décimo de segundo durante o qual a ação corretiva é tomada, durante o qual alguma mudança é feita nas características da resposta (*output*) manual do homem. Segue-se um período de cerca de dois décimos de segundo no qual nada de novo é feito, e então outro período de cerca de um décimo de segundo no qual novas correções são feitas, e assim por diante (MACY 6, 1950, p. 33, tradução nossa).

Como na primeira conferência, os demais membros também apresentaram suas considerações sobre a teoria da percepção de Stroud, que destacava o papel importante da descontinuidade no processamento de informação. Alguns participantes desse debate buscaram explicar essa descontinuidade do processo perceptivo a partir do funcionamento bioinformático do organismo. Gerard, por exemplo, defendeu que o caráter projetivo da ação humana seria não apenas psicológico, mas também fisiológico, enquanto para Wiener a descontinuidade observada na percepção é consequência direta do processamento digital de informações pelo cérebro, isto é, do fato de os neurônios operarem como interruptores. O diálogo entre estes dois pode ser encontrado em Macy 6 (1950, p. 33-34, tradução nossa):

> *Gerard*: Sobre esse assunto, imagino se não é o mesmo fenômeno que se encontra em muitos estudos do sistema nervoso. Se você pegar um sujeito completamente destreinado, sentá-lo numa cadeira, provocar o movimento espasmódico do joelho a intervalos regulares e então parar de martelar [o joelho] sem avisar, a perna provavelmente continuará a chutar por várias vezes, nos intervalos de tempo "esperados" [...]
>
> *Wiener*: Posso dizer por que deve ser este o caso? A transmissão [de informações no organismo] é essencialmente descontínua. Seus neurônios individuais ligam ou desligam. É apenas por meio de amostragem que se obtém algo próximo de uma entrada contínua. Se pegarmos um tempo muito breve sua amostragem será decididamente ruim. A média de informações de entrada chegando não será atingida de nenhuma forma precisa. Portanto, para ter

> uma correção realmente significativa a ser feita, você deve esperar. Penso que esta é uma das principais razões de por que você tem essa performance discreta [da percepção].

Ainda na sexta conferência Macy, ocorreram outras discussões, fruto desse mesmo espírito do movimento cibernético. A terceira discussão do dia foi feita pelo psicanalista Lawrence Kubie, referente às dificuldades postas pelas neuroses à adaptação da vida em sociedade. Com o título "Neurotic Potencial and Human Adaptation", Kubie, em Macy 6 (1950), abordou a seguinte questão: é verdade que o princípio biológico de adaptação ao meio é uma "lei natural"? Ou seja, seria o comportamento socialmente valorizado um demonstrativo de uma bem-sucedida adaptação ao meio? Imaginando que essa característica o fosse, como pensar os frequentes casos de indivíduos plenamente adaptados que sofrem devido a profundas neuroses? Ou os casos de indivíduos totalmente adaptados a um certo tipo de ambiente, contudo, completamente inadaptados a outros? Seria a neurose sinônimo de comportamento inadaptado? Ou o preço da adaptação seria o desenvolvimento de uma neurose? Kubie, em Macy 6 (1950), tratou desse problema em termos cibernéticos de *feedback* e de análises clínicas.

Na última conferência do dia, o físico alemão Heinz von Förster efetuou uma discussão envolvendo um modelo para a memória. Nessa apresentação, intitulada "Quantum Mechanical Theory of Memory", "todos" os aspectos da memória foram modelados em elementos discretos, como a sua dimensão fenomenológica, psicológica e mesmo biofísica. Von Förster, em Macy 6 (1950), divide sua exposição conforme esses três aspectos, abordados sucessivamente. Inicia pela questão da relação entre o tempo físico e o tempo psicológico.

Conforme Von Förster, em Macy 6 (1950, p. 112, tradução nossa), "[...] nossa memória funcionaria como um gravador: qualquer informação recebida seria armazenada indefinidamente" e sua recordação traria o evento percebido de volta, em sua integridade, inclusive de sua duração temporal. Como não é isso que se verifica fenomenologicamente, ao se estudar ou experienciar a memória, é certo que, enquanto o tempo passa, perdemos certa quantidade de informação por esquecimento. Enfatiza Von Förster, em Macy 6 (1950, p. 112, tradução nossa):

> A ideia principal é que todo evento observado deixa uma impressão que pode ser dividida em várias impressões elementares. Penso que é justificado assumir isso porque os órgãos dos sentidos também estão divididos em vários receptores sensórios elementares.

O conteúdo da memória estaria dividido em diversas impressões elementares, oriundas dos discretos receptores sensórios do corpo. Com o tempo e a ação do esquecimento, esse número de impressões elementares armazenadas na memória se modifica. Von Förster procurou, por conseguinte, uma função matemática capaz de relacionar essas duas quantidades de elementos. Por analogia com os fenômenos físicos e químicos de decaimento, Von Förster, em Macy 6 (1950, p. 112, tradução nossa), assumiu que "[...] a taxa de variação do número de impressões elementares existentes por unidade de tempo deve ser proporcional ao número de impressões elementares existentes", o que conduz a uma certa função determinada por uma constante, um coeficiente de esquecimento. Quanto maior o coeficiente de esquecimento, mais rápido o esquecimento de certo conteúdo de memória, e vice-versa. Tal função reside não somente na analogia com fenômenos físicos de decaimento e desorganização já conhecidos, mas sobretudo na proximidade existente entre o que ela prevê e os resultados obtidos em experimentos sobre a memória.

O segundo dia da sexta conferência Macy teve início com uma discussão conjunta sobre os mecanismos de recordação e reconhecimento de informações armazenadas na memória. Intitulada "Possible Mechanisms of Recall and Recognition", a discussão começou com considerações de diversos participantes sobre a necessidade da comunicação interdisciplinar, especialmente entre as ciências duras, naturais, e as ciências humanas. Isso pode ser constatado no relato de Heinz von Förster em Macy 6 (1950, p. 148, tradução nossa): "Temos que aprender a falar em algum nível de ciência comum [...] começar de algo sobre o qual o cientista social e o físico possam dizer 'Sim, concordamos sobre isto. Estamos falando a mesma linguagem. Como podemos progredir a partir daqui?'"

Irritados com a abordagem feita pelos demais conceitos mecanicistas de certos participantes e com o que lhes pareceu falta de disposição ou capacidade de escutar, alguns psicólogos e cientistas sociais apelaram a seus colegas para que se preocupassem com os próprios pontos

cegos num grupo que realmente tinha como foco a comunicação entre pessoas. Esse pequeno "puxão de orelhas", de fato, surtiu efeito, pois a discussão propriamente científica que se seguiu atentou, bem mais que no dia anterior, para a lógica própria das dimensões sociais e psicológicas dos fenômenos abordados (memória, percepção, neurose).

Nenhuma conceituação realmente nova foi feita, de modo que a discussão toda consistiu num refinamento dos tópicos abordados no dia anterior, buscando delinear melhor o funcionamento dos mecanismos da memória, visando a identificar os processos específicos de recordação e reconhecimento de informações/conteúdos. Diversas conceituações, experimentos e hipóteses explicativas foram levantadas, numa dinâmica nem sempre linear de discussões; alguns assuntos foram retomados diversas vezes, outros atropelados por novos temas. Entretanto, pode-se dividir essa discussão em dois grandes blocos temáticos: o dos mecanismos analógicos de acesso à memória e o dos mecanismos digitais.

Segundo os relatos de Heinz von Förster em Macy 6 (1950), o debate revelou duas características principais a propósito dos mecanismos de acesso à memória, discutidos por meio de categorias cibernéticas, em especial a analógica e a digital. A memória, por exemplo, foi entendida como corporificada sobretudo por circuitos digitais reverberantes, isto é, circuitos sujeitos a uma causalidade circular, em que a informação de saída reentra e refaz o mesmo caminho, reiteradamente. O acesso a esses circuitos lacrados aparentemente só pode ser feito por mecanismos analógicos, os quais consistem em mensagens buscadoras, que se difundem de modo contínuo pelo organismo até encontrar alguma rede que se interesse por elas, ou de alterações analógicas dos níveis de limiar sináptico, os quais possibilitam a captura dos nós de redes sinápticas lacradas em circuitos reverberantes por outras redes. Nenhum mecanismo propriamente digital de acesso a essas memórias parece poder ser desfeito. Em segundo lugar, tal discussão explicitou um pressuposto que já vinha operando nas discussões anteriores, a unidade entre físico e psíquico, baseada na analogia de mecanismo entre os processos de cada domínio. Mecanismos de *feedback* foram atribuídos a interações biofísicas e psicológicas; dentre as várias ocorrências, a mais interessante é a teoria de Frank, para a qual um organismo humano aprende padrões de comportamento novos, ao selecionar, dentre as várias mensagens internas e externas a que está continuamente exposto, aquelas

que lhe interessam e as utiliza como dados de entrada nos diversos circuitos de retroalimentação que o compõem. De fato, funcionando segundo o mesmo mecanismo de *feedback*, os níveis psicológico e biofísico se comunicam, o que permite vê-los como unos, no que tange à ação de causas. Dessa forma, um evento propriamente psicológico, como a atribuição de significado a uma situação, pode causar alterações que são simultaneamente de personalidade (mudanças no significado atribuído a mensagens externas e padrões de resposta apreendidos) e fisiológicas (alterações na seletividade e sensitividade de determinados sistemas de órgãos). A conceptualização cibernética funciona, ao mesmo tempo, como veículo de conceituação e garantia de unidade (por analogia de mecanismo) entre os níveis biofísico e psicológico-social.

Exposto brevemente por Wiener, o último assunto abordado na sexta conferência, intitulada "Sensory Prosthesises", diz respeito à substituição de receptores sensórios humanos por máquinas protéticas. Wiener relatou a construção de uma máquina para substituir a audição dos completamente surdos. Trata-se de uma máquina transdutora, a qual converte as vibrações sonoras em vibrações mecânicas sentidas pelos dedos. O padrão vibratório do som é um pouco alterado, para garantir maior discernimento e inteligibilidade. Segundo Wiener em Macy 6 (1950, p. 204, tradução nossa):

> (a) palavras distinguíveis são reconhecidas como diferentes [...] em vários casos aprendemos a fazer pessoas reconhecerem um pequeno vocabulário [...] (b) as mesmas palavras ditas por pessoas diferentes são reconhecíveis como possuindo o mesmo padrão.

Wiener, em Macy 6 (1950), alude também a uma máquina para auxiliar a locomoção de cegos, projetada para dar-lhes indicações sobre os contornos dos objetos presentes no ambiente. Trata-se de duas células fotossensíveis conectadas ao ouvido e programadas para identificar o deslocamento espacial. As células funcionam ao serem apertadas pela mão, quando cada uma emite um som diferente. Porém, quando próximas de algum objeto, os dois sons coincidem. Assim, o cego é capaz de saber a que distância se encontra dos objetos do ambiente pelo grau de similaridade entre os sons emitidos pelas duas células fotossensíveis.

Supõe-se que houve, no binômio humano-máquina, duas operações cibernéticas atuantes. Uma, a tradução de informações sensórias em informações de outro tipo – conversão de som em movimento táctil, conversão de deslocamento da distribuição luminosa ambiente em vibração sonora. A informação transformada serve de substituto para a informação captada; esta muda de meio, no entanto, conserva uma forma equivalente, mantendo, em outro suporte, a referência ao mundo exterior. Outra operação cibernética é a existente entre um humano acoplado à máquina e o ambiente externo: as informações colhidas pela máquina possibilitam ao ser humano estar em contínuo *feedback* com o mundo à sua volta, orientando suas ações pelas informações substitutas fornecidas pela máquina.

A sétima conferência Macy (MACY 7, 1951), ocorrida em 23 e 24 de março de 1950, foi dedicada à questão da linguagem. Esta foi pensada como um código submetido a processos de codificação e transmissão de informações (fala, escrita), recodificação (tradução) e transcodificação (processos sociais de mudança linguística). Esses processos atuam simultaneamente em vários níveis: no nível biofísico (cérebro é computador com código lógico-digital), no nível psicoperceptivo (produção, recepção e compreensão da língua são operações de codificação, tradução e decifração), no nível comportamental (comportamento neurótico é déficit de simbolização pela linguagem) e no nível "ideal" ou estrutural (língua enquanto estrutura abstrata de relação entre elementos discretos).

Conforme os relatos de Heinz von Förster em Macy 7 (1951), a sétima conferência Macy se iniciou com a apresentação do fisiologista Ralph Gerald, tratando da natureza dos mecanismos do sistema nervoso, levando em consideração se eles são analógicos ou digitais. A segunda sessão do primeiro dia contou com a exposição do psicólogo experimental J.R. Licklider, que abordou a capacidade humana de reconhecimento de falas distorcidas ou alteradas. A quarta apresentação foi do engenheiro Claude Shannon, que teve como objetivo discutir a redundância da escrita em inglês. O segundo dia da sétima conferência Macy começou com um especial interesse na relação entre a noção de código e sentido. A primeira apresentação do dia foi feita por Margareth Mead, sobre como aprender línguas "primitivas", usualmente dotadas de estruturas gramaticais muito exóticas para os padrões ocidentais. Em seguida, os

psicólogos Heinz Werner e John Stroud falaram sobre o problema do desenvolvimento ontogenético da linguagem, em especial os processos de tomada de consciência semântica e sintática durante a infância. O ciclo de discussões se encerrou com a apresentação de Lawrence Kubie, que propôs teorizar as diferenças fundamentais entre a neurose e a normalidade. O objetivo dessa última apresentação consistia em saber até que ponto certos comportamentos não linguísticos poderiam ser compreendidos sob o ponto de vista da linguagem pensada como código.

A oitava conferência Macy (MACY, 1952), realizada nos dias 15 e 16 de março de 1951, teve como tema a comunicação entre os seres humanos. Principiou com a fala do psicólogo Alex Bavelas, sobre suas experimentações com a dinâmica entre pessoas organizadas em grupo. O objetivo desse estudo era averiguar as implicações do formato dos canais de comunicação interpessoais, classificando, assim, a eficácia na realização de tarefas. A segunda sessão foi aberta pela fala do pedagogo Ivo Richards, na qual se dedicou a saber até que ponto é válido ou mesmo possível pensar a linguagem, sendo que o único instrumento disponível para tal tarefa é a própria linguagem. A terceira apresentação ficou a cargo do psicanalista Lawrence Kubie, com o tema da comunicação simbólica humana em diferentes estados de consciência, bem como da distinção entre a comunicação consciente e a comunicação inconsciente. O segundo dia teve início com a apresentação do estudioso de psicologia animal Herbert G. Birch, em que se levantou a questão da existência da comunicação complexa entre animais.

A nona conferência Macy (MACY 9, 1953), promovida nos dias 21 e 22 de março de 1952, teve dois temas principais. O primeiro, contrariando em alguma medida o que dissemos no começo deste capítulo sobre as características da ciência cognitiva, são as emoções humanas: código particular, que permite a presença de emoções no processo de comunicação humana. O segundo diz respeito aos processos orgânicos de homeostase e o de aprendizado de novos padrões comportamentais. A nona conferência se inicia com a fala de Bateson sobre as dinâmicas comunicativas responsáveis pelo riso e emoções similares na comunicação humana. A segunda discussão foi conduzida por Kubie, acerca dos tipos de relação causal existentes entre as emoções e o comportamento. A terceira foi a propósito da obra de Ross Ashby, em que foram expostos os

princípios de uma máquina criada por ele, o homeostato, desenhada para ser um modelo de dinâmica de solução para problemas adaptativos dos organismos e descobrir, nesse sentido, como os organismos aprendem e conseguem reorganizar seu equipamento neurológico cerebral para que, por mais incomum que seja um meio ambiente, ele possa aprender e tomar uma ação apropriada.

Por fim, a décima e última conferência Macy (MACY 10, 1955) aconteceu nos dias 22, 23 e 24 de abril de 1953, em Princeton. Essa conferência contou com três comunicações: uma do neurologista britânico Grey Walter, sobre o cérebro; outra, do filósofo israelense Yehoshua Bar-Hillel, a respeito da possibilidade de se medir a informação semântica; e a última, do linguista sino-americano Yuen Ren Chao, sobre os mecanismos responsáveis pelo sentido, na língua. A décima conferencia Macy representava o fim de um ciclo de conferências e do grupo dos cibernéticos.

Quanto aos participantes das Conferências Macy, conforme podemos notar, a partir dos temas e problemas levantados anteriormente e mais detalhadamente abordados na sexta conferência Macy, destaca-se uma grande interdisciplinaridade, um vocabulário calcado em conceitos cibernéticos e uma diversidade de temas. Nesse primeiro momento da cibernética houve a tentativa de se estabelecer uma ciência fisicalista, centrada em uma nova concepção de máquina, que introduziu a conceituação lógico-matemática de processamento de informação que serviu como fundamento para a Inteligência Artificial, a ser tratada no próximo capítulo. Antes de fazermos uma análise deste momento inicial da ciência cognitiva, com base nos encontros relatados acima, falamos, a seguir, do surgimento desta área de pesquisa propriamente dita.

2.4 O contexto da cibernética e o surgimento da ciência cognitiva

De acordo com Varela (1991), um cientista cognitivo da geração mais nova e dissidente da velha guarda, a fase da cibernética propiciou inúmeros resultados teóricos e práticos que influenciaram o desenvolvimento posterior da ciência cognitiva, em especial, o surgimento da teoria da informação, apresentada como uma teoria estatística dos sinais e canais de comunicação. Nas últimas conferências Macy pode-se

perceber a presença de pesquisadores fundadores da Teoria Matemática da Comunicação, como Shannon, por exemplo, na sétima conferência Macy. Os primeiros exemplos de robôs parcialmente autônomos, capazes de incorporar ou pelo menos visualizar uma possibilidade de uma auto-organização parcial; a escolha da análise lógico-matemática para descrição do funcionamento do sistema nervoso; e a instauração da teoria dos sistemas, como uma disciplina capaz de formular as regras e princípios gerais de qualquer sistema complexo.

Parte destes inúmeros avanços estão interligados ao fato de que a cibernética tem surgimento atrelado à proposta de um agrupamento de diversos estudos pertencentes a outros campos, como, por exemplo, a matemática dos fenômenos aleatórios, a análise e correções de desvios e ruídos em redes de transmissão de sinais, os tópicos em engenharia de dispositivos automáticos e semiautomáticos, os tópicos em engenharia elétrica e de instrumentos de comunicação, a teoria formal dos autônomos, os estudos das bases neurofisiológicas do comportamento humano e animal, os estudos sobre a tomada de decisões dos agentes sociais, com base na informação disponível, os estudos sobre o equilíbrio em organismos e em comunidades, inclusive humanas, e o projeto de máquinas de calcular ultrarrápidas (computadores). Embora cada tema possua uma evidente complexidade própria, esses estudos possuem um ponto convergente em comum, em especial o tratamento dos diversos problemas de controle, quer em sistemas artificiais, quer em processos naturais.

Os problemas de controle estão ligados a uma certa relação de causalidade circular. Em geral, os sistemas variam os seus estados, levando em conta parte dos resultados de suas ações anteriores, oferecendo dados que "realimentam" (*feedback*) o sistema. Wiener (1961) teve algum contato com praticamente todos os campos acima citados. Ele notou que as mesmas reflexões sobre o controle e a causalidade circular também apareciam, tanto nos problemas ligados as áreas da engenharia quanto em biologia, e que essa coincidência estava relacionada a importantes questões metodológicas da área.

Ademais, Wiener (1961) e seus colaboradores, na tentativa de desenvolver os princípios de sua área, propuseram uma perspectiva e uma taxonomia amparada no comportamento observável dos sistemas. Quando esse comportamento revela algum tipo de relação de *feedback*, utiliza-se o

termo teleológico, no sentido de representar o controle por meio de *feedback*, em vista de um objetivo. O termo teleológico aparece na cibernética com uma definição diferente da habitualmente usada na filosofia. Imagina-se que sistemas observados pelo prisma teleológico da cibernética podem ser seres vivos ou máquinas, os quais, sob essa perspectiva, podem ser estudados sem que seja necessário considerar as suas diferenças estruturais.

O conceito de teleologia oferecido por Wiener, Rosenblueth e Bigelow toca em um tema consagrado e polêmico da filosofia. Após algumas revisões, os autores insistiram em classificar o termo teleologia como sinônimo de "propósito controlado por *feedback*", acrescentando:

> No passado, a teleologia foi interpretada como implicando propósito, acrescentando-se muitas vezes o vago conceito de uma "causa final". Essa ideia de causas finais levou à oposição entre teleologia e determinismo. Uma discussão sobre causalidade, determinismo e causas finais está além do escopo deste ensaio. Pode-se ressaltar, no entanto, que o caráter de ser proposital (*purposefulness*), tal como definido aqui, independe bastante da causalidade inicial ou final. A teleologia foi desacreditada principalmente por ter sido definida como implicando uma causa subsequente no tempo a um efeito dado. Ao se rejeitar esse aspecto da teleologia, contudo, rejeitou-se também, infelizmente, o reconhecimento, que lhe é associado, da importância do propósito.
>
> Restringimos a conotação de comportamento teleológico aplicando essa designação apenas a reações propositadas que são controladas pelo erro da reação – isto é, pela diferença entre o estado, num dado momento, do objeto que se comporta, e o estado final, interpretado como propósito. Comportamento teleológico torna-se assim sinônimo de comportamento controlado por realimentação negativa; ganhando, portanto, em precisão, por ser uma conotação suficientemente restrita.
>
> De acordo com esta definição limitada, a teleologia não se opõe ao determinismo, mas à não teleologia. Tanto os sistemas teleológicos como os não teleológicos são determinísticos quando o comportamento considerado pertence ao domínio ao qual o determinismo se aplica. O conceito de teleologia tem apenas uma coisa em comum com o conceito de causalidade: um eixo temporal. A causalidade, contudo, implica uma relação funcional unidirecional e relativamente irreversível, enquanto o que concerne

> à teleologia é o comportamento, e não as relações funcionais (WIENER; ROSENBLUETH; BIGELOW, 1943, p. 49-50, tradução nossa).

O caráter dinâmico da definição dos sistemas estudados pela cibernética enseja certa comparação entre os organismos biológicos e os artefatos construídos pelo ser humano. Wiener (1961) faz um uso forte e sistemático da analogia entre máquinas e animais, empregando o termo autônomo em referência a ambos.

O conceito de autômatos não surge como um simples recurso retórico: acima de tudo, é um método de abordagem, capaz de evidenciar a natureza do assunto, como mostra a seguinte passagem de Wiener (1961, p. 70) sobre os autômatos:

> O estudo mais recente dos autômatos, sejam de metal ou de carne e osso, é um ramo da tecnologia da comunicação: suas noções cardeais são as de mensagem, quantidade de distúrbio ou "ruído" – termo tomado da tecnologia telefônica –, quantidade de informação, técnica de codificação, e assim por diante.
>
> Numa teoria desse tipo, lidamos com autômatos efetivamente ligados ao mundo exterior, não apenas por seu fluxo de energia, seu metabolismo, mas também por um fluxo de impressões, de mensagens que chegam, e das ações e mensagens que saem. Os órgãos pelos quais as impressões são recebidas são os equivalentes dos órgãos sensoriais animais e humanos. Compreendem as células fotoelétricas e outros receptores de luz; sistemas de radares, que recebem suas próprias ondas curtas hertzianas; registros de potencial de hidrogênio, que podem ser chamados provadores; termômetros; medidores de pressão de vários tipos; microfones; e assim por diante. Os efetuadores podem ser motores elétricos ou solenóides ou serpentinas de calefação ou outros instrumentos de espécies muito diversas.

O método implicado na proposta de Wiener (1961), o qual, por ora, chamaremos de "método cibernético", é uma variante ou parte do chamado "método sistêmico": ambos admitem que a determinação do real

não deriva apenas dos seus elementos constitutivos, mas também de sua estrutura relacional e funcional.

O "método sistêmico" é um esforço por conseguir a visão mais completa possível do objeto estudado, mediante a inclusão de um número cada vez maior de elementos, partes ou relações. Cada parte do todo é considerada, em princípio, como sendo relevante, assim como as relações entre as partes e as relações entre o objeto e seu meio circundante. Essa inclusão se dá sob a forma de uma integração: um tipo de descrição estrutural-relacional-funcional que torna visíveis os vínculos que formam o objeto.

O método sistêmico se contrapõe ao chamado "método analítico". Se o método analítico visa a identificar fatores causais relevantes isolados, para depois montar uma "causalidade composta", o método sistêmico, pelo contrário, aponta para uma "causalidade coordenada", a qual emerge da própria convergência dos elementos considerados.

Por um lado, as discussões expostas neste capítulo mostram a tentativa, pelo menos de alguns participantes do grupo de pesquisadores, de criar uma teoria do controle, da possibilidade de manipular, dominar, administrar certos fenômenos ou funcionamento de certos sistemas. No entanto, em outra frente de discussões e propostas metodológicas, embora ainda insipiente, parte da abordagem subjacente a esta perspectiva consiste em adotar um método baseado na complexidade, em oposição ao método analítico, tradicionalmente defendido na posição cartesiana. Além de uma mudança nas bases metodológicas de pesquisa, parte dos integrantes dos encontros relatados anteriormente propõe, em maior ou menor grau, rearticular a fragmentação das áreas científicas estabelecida na prevalência do método analítico. Neste cenário, parte do grupo de pesquisadores externalizou o pressuposto da presença da desordem no mundo e, como consequência, a incerteza tanto no tocante aos fenômenos quanto no tocante à sua possibilidade de conhecimento. Nesse sentido, ao pressupor uma nova concepção de mundo, tal grupo de pesquisadores alinha-se também a uma nova concepção de ciência.

Dentre os primeiros difusores do que seria posteriormente a ciência da complexidade encontramos o coautor, juntamente com Claude Shannon, citado anteriormente, de *The Mathematical Theory of Communication*, Warren Weaver (1948). Ele estabelece três tipos de problemas relativos ao

contexto científico da complexidade: da simplicidade, da complexidade desorganizada e o problema da complexidade organizada.

O primeiro deles trata das questões que envolvem um número limitado e baixo de variáveis, cuja análise pode ser realizada através do emprego de equações. Este seria o caso, por exemplo, da trajetória de corpos em movimento linear em ambientes controlados. O problema de complexidade desorganizada envolve uma quantidade grande de variáveis, que fogem a uma análise em que seja possível controlar as variações no sistema. Esse tipo de problema é analisado por metodologias de estatística e probabilidade, como, por exemplo, das leis de pressão e temperatura atmosférica. Por fim, o problema da complexidade organizada possui um número de variáveis que se comunicam de forma inter-relacionada, e suas interações ocorrem em um todo orgânico, de modo não linear. Este seria o caso do estudo sobre a cognição, especialmente humana, considerando que os processos cognitivos ocorrem em sistemas complexos, especialmente se considerarmos que a base física em que eles são instanciados (no caso do cognitivismo, a ser apresentado no próximo capítulo) ou realizados (no caso do conexionismo, também exposto no próximo capítulo) é o cérebro.

A peculiaridade do método cibernético, nesse contexto, é a sua preferência por enfocar a evolução temporal dos objetos estudados, que são vistos como processos dotados de certa persistência ou direcionamento. Mais do que a "causa do objeto", o que se procura identificar e descrever são as "causas da estabilidade" – tanto do objeto em suas transformações como do seu modo de proceder.

À primeira vista, a descrição dos fenômenos físicos consiste em mostrar como eles se enquadram nas chamadas "leis do movimento", que, por sua vez, só podem ser compreendidas no marco de certo número de "leis de conservação". Quando o movimento de um sistema não pode ser descrito em função de certos parâmetros constantes, diz-se haver "instabilidade". A descrição de sistemas "instáveis" só é possível quando as "fontes" dessa instabilidade obedecem igualmente a certos requisitos, podendo ser tratadas como introdutoras de alterações "suaves" ou "previsíveis" nos parâmetros conservativos do sistema original. Mesmo assim, a instabilidade favorece a presença de imperfeições na descrição.

Em função da sua "capacidade de previsão", quando o estado futuro de um sistema não é de alguma forma "calculável" (ou pelo menos

fixada a sua probabilidade), tendo em vista o seu passado, a descrição falha, e será preciso reformular a abordagem. No caso de sistemas em que a complexidade é muito elevada, seja na quantidade de elementos, seja nas relações entre eles, seja das variáveis envolvidas, como em fenômenos biológicos, certas abordagens são tidas como impossíveis ou são tratadas em termos de probabilidade.

Segundo Dupuy (1996), o pensamento cibernético é oriundo de toda uma evolução no modo de pensar a representação do conhecimento, possibilitando que o modelo de representação científica fosse estendido ao domínio dos problemas da mente. A ciência ocidental é marcada, nas palavras de Dupuy (1996, p. 21), pela ideia de que "[...] só podemos conhecer aquilo de que somos causa, o que fabricamos", ou seja, só podemos conhecer a partir da imitação da natureza e pela representação, através da produção de experimentos. Dessa maneira, graças ao recurso da analogia e ao processo de conhecimento, por intermédio da representação científica, poderiam ser reproduzidos aparatos técnicos mecanicamente capazes de representar o próprio processo do conhecimento, tal como a máquina de Turing ou as redes neurais artificiais, a serem discutidos no próximo capítulo, tomados como modelos adequados para a explicação de certos processos cognitivos. No caso do cognitivismo, para Dupuy (1996, p. 36), as descobertas lógicas de Gödel e de Turing teriam ajudado a conduzir a esta ideia: "O pensamento, essa atividade psíquica, essa faculdade do espírito que tem o conhecimento como objeto, nada mais é, afinal, do que um processo mecânico ordenado, um automatismo 'cego'".

O pensamento cibernético teria sido composto com base em uma evolução no pensamento científico. Supondo-se a ideia de Turing (1950) de que pensar é calcular através da produção de um modelo do objeto conhecido, também seria possível imaginar um modelo capaz de produzir e representar um modelo de funcionamento de um sistema capaz de conhecer, ou seja, uma representação da faculdade de representação. Nessa perspectiva, o pensamento cibernético compreendia o ato de conhecer, a partir da simulação, isto é, como a capacidade de modelagem que consiste em reproduzir o funcionamento de um sistema.

O pensamento contemporâneo ao movimento cibernético previa-lhe um futuro promissor. A cibernética estimulava a imaginação de toda uma época, pois havia nascido como um sucesso midiático,

amplamente divulgada pelos meios de comunicação de grandes potências econômicas como os Estados Unidos e países da Europa, anunciando uma vasta quantidade de previsões e transformações sociais e tecnológicas. As mesmas revistas que em 1948 apregoavam que a cibernética crescia de maneira descontrolada e problemática, poucos anos depois afirmavam:

> Três anos atrás Norbert Wiener, professor de matemática no M.I.T., era um "cabeludo" que havia cunhado a palavra "cibernética" para envolver os muitos lados da ciência dos dispositivos de comunicação e controle. Agora o livro de Wiener, *Cibernética*, é um clássico, e Wiener é um profeta que é ouvido por sisudos homens de negócio de cabelo curto. Muitos deles concordam sinceramente que a "revolução cibernética" que ele predisse já está em progresso (TIME MAGAZINE, 1950, *apud* MASSARO, 2010, p. 19).

O movimento cibernético, com a sua estreia espetacular, logo deu início a um processo de institucionalização. Com o avançar da década de 1950, cursos universitários de cibernética começaram a surgir ao redor do mundo, especialmente na Europa. Embora, em sua terra natal, a cibernética jamais tivesse alcançado o título de disciplina científica, pouco a pouco, após os anos 1950 e até o início da década de 1970, começaram a aparecer manuais de cibernética, a fim de difundir e popularizar a área, pelo menos no meio acadêmico. Dentre tais livros introdutórios, citamos os seguintes, com seu ano de publicação:

1956 – *An Introduction to Cybernetics* – W. Ross Ashby

1958 – *Cibernética* Técnica (em russo) – L.P. Kraizmer

1959 – *Cybernetics and Management* – Stafford Beer

1960 – *Cybernetics Without Mathematics* – Henryk Greniewski

1961 – *An Approach to Cybernetics* – Gordon Pask

1964 – *Vvedeniye v Kibernetiku* (Introdução à cibernética) – Viktor Glushkov

1965 – *Cybernetics and Biology* – F.H. George

1967 – *Cybernétique et Biologie* – Andrée Goudot-Perrot

Todos esses manuais tinham em comum a ideia de que a cibernética seria uma nova ciência, com objetivos próprios, modos de quantificação e conceitos próprios. Esses manuais possuem a intenção de apresentar os fundamentos de uma ciência específica e resumir o estágio atual de seus conhecimentos, oferecendo, assim, tudo de que um cientista precisaria para começar a trabalhar.

Para um observador da época, tudo indicava que o futuro pertencia à cibernética. Entretanto, pouco a pouco, passou-se a falar cada vez menos desse movimento: talvez estivessem todos preocupados com os movimentos sociais e históricos da época ou talvez a revolução cibernética tivesse demorado demais; o fato é que, no final da década de 1970, a nova ciência evocava um ar antiquado. Nesse período de 1970-1980, surgiam inúmeras novas estrelas no céu da ciência: teoria dos sistemas, Inteligência Artificial, teoria da complexidade, ciência cognitiva, auto-organização, autopoiésis, dentre outras. Somos tentados a pensar que a história desse período é marcada por revoluções a cada minuto. Em meio a tantas estrelas, talvez a cibernética tenha caído no esquecimento e tenha sido apenas uma estrela ofuscada pelo brilho de tantas outras.

Como podemos perceber, e voltaremos a tratar no próximo capítulo, de alguma maneira, o que chamamos de ciência cognitiva é proveniente do movimento da cibernética. No entanto, o suposto parentesco que propomos sofre uma resistência, por parte dos descendentes, em admiti-lo. A ciência cognitiva parece possuir certa aversão à cibernética. Em geral, cientistas cognitivos entendem que sua proposta é epistemológica, visando conhecer certos fenômenos do mundo, sem o viés prático ou com fins de domínio da natureza. No entanto, poderíamos dizer que os encontros e embates registrados nas conferências expostas neste capítulo estabeleceram as bases para o surgimento da ciência cognitiva tal como instituída posteriormente, seja ela entendida como uma nova área independente, seja como uma dissidência, seja como aperfeiçoamento da cibernética. Nesse sentido, a ciência cognitiva não veio do nada, mas teve uma origem, cujas bases foram estabelecidas, de uma forma ou de outra, sob influência dessa área da física já previamente estabelecida.

Em resumo, apesar de ainda aparentemente não constituir um paradigma consolidado, o grupo de pesquisadores com suas pesquisas, hipóteses, teorias, seus encontros e atuação no sentido de divulgar

os trabalhos e desafiar novos adeptos propiciaram o surgimento da ciência cognitiva, favorecendo as condições tanto epistemológicas quanto heurísticas e financeiras para o seu desenvolvimento, como discutiremos a seguir.

3

A ciência cognitiva em sua fase inicial: contexto epistêmico

Apresentação

Neste terceiro capítulo do livro expomos as bases teóricas e epistemológicas que habitavam a ciência cognitiva em sua fase inicial. Para isso, dividimos o capítulo em quatro seções. Na primeira seção apresentamos os primeiros anos da ciência cognitiva, após o fim da cibernética, tomando como referência o Relatório do Estado de Arte solicitado pela Fundação Sloan. Na segunda seção, a fim de evidenciar as suas bases metodológicas, expomos alguns conceitos básicos dessa área, principalmente os de modelo e representação, buscando, como em todo o texto, expor razões para o estabelecimento de seu estatuto científico. Baseados neste intuito, nas duas **últimas** seções apresentamos duas das principais perspectivas daquele momento que poderiam se configurar como paradigmas dominantes. A primeira delas é o cognitivismo, tratado na terceira seção. Também chamada de Inteligência Artificial, em seu sentido mais forte, essa vertente conjectura que a cognição funciona de maneira idêntica ou semelhante aos procedimentos computacionais, podendo ser modelada por máquinas do tipo Turing. Na quarta seção examinamos as bases de outra grande concepção, o conexionismo, também chamado de Redes Neurais Artificiais. Essa vertente compreende que a cognição é resultado do processamento, distribuído e em paralelo, feito pelas várias unidades simples que compõem uma rede neural. Seriam essas posições

duas candidatas a paradigmas, seriam teorias distintas de um paradigma apenas ou nenhuma dessas alternativas? Esperamos estabelecer as bases desta discussão que será realizada nas considerações finais.

3.1 A ciência cognitiva após a Cibernética

No capítulo anterior procuramos mostrar alguns elementos históricos que propiciaram o surgimento da ciência cognitiva, visando refletir sobre seu estatuto científico nesta fase. Entretanto, como também procuramos mostrar, ela ainda não havia se estabelecido como uma área de pesquisa formalmente constituída. Ainda que em gérmen, os pesquisadores trabalhando em torno do problema da inteligência ou dos processos cognitivos, mais genericamente falando, foram dando forma a essa nova área de pesquisa incipiente.

Existe um consenso na história da ciência cognitiva de que esta foi reconhecida oficialmente em 1956. De acordo com Miller (1979), a comunidade de cientistas cognitivos ganhou vida no Simpósio sobre Teoria da Informação realizado no MIT, de 10 a 12 de setembro de 1956. O segundo dia desse encontro se destaca, para Miller (1979), graças à exposição de alguns trabalhos de grande impacto. Um deles, apresentado por Allen Newell e Herbert Simon, intitulado "Logic Theory Machine", focalizava a primeira prova completa de um teorema executado em uma máquina computadora. O segundo trabalho de grande impacto foi apresentado pelo linguista Noam Chomsky, intitulado "Three Models of Language". Neste texto, Chomsky mostra que um modelo de linguagem derivado da teoria da informação, proposto por Shannon e Weaver (1949), não poderia, de forma alguma, ser aplicado com êxito à "linguagem natural". Assinala Gardner (1996, p. 44) sobre as considerações de Miller:

> Saí do simpósio com uma forte convicção, mais intuitiva que racional, de que a psicologia experimental humana, a linguística teórica e a simulação computacional de processos cognitivos eram todas partes de um todo maior, e de que o futuro veria uma crescente elaboração e coordenação de seus interesses comuns [...] Eu venho trabalhando por uma ciência cognitiva há aproximadamente vinte anos, tendo começado antes de saber como chamá-la.

Como já assinalamos no capítulo anterior, em meados da década de 1950 os neurocientistas estavam começando a registrar impulsos de neurônios individuais do sistema nervoso. No MIT, a equipe de pesquisa de McCulloch, dirigida pelos neurofisiologistas Jerome Lettvin e Humberto Maturana, havia feito um registro do funcionamento da retina de uma rã. Eles haviam conseguido mostrar que os neurônios são sensíveis a formas extremamente específicas de informação, como, por exemplo, pequenos pontos escuros semelhantes a insetos, os quais se moviam através de seu campo perceptivo.

Dentre outras descobertas nos ramos da antropologia e da neurociência, ainda em 1956, um grupo de cientistas, com formação em matemática e lógica e interessados nos problemas dos computadores, reuniu-se no Dartmouth College para discutir seus trabalhos. Nessa faculdade, estava concentrada a maior parte dos cientistas desenvolvendo suas investigações com o que viria a ser denominado "Inteligência Artificial" (IA). Nesse grupo estavam, inclusive, os que geralmente são considerados os pais fundadores da IA: John McCarthy, Marvin Minsky, Allen Newell, Noam Chomsky e Herbert Simon. Durante o encontro, foram examinadas ideias para programas que iriam solucionar problemas, reconhecer padrões, raciocinar logicamente, tendo sido determinadas as principais questões a serem debatidas nos anos seguintes. A despeito de não ter emergido qualquer síntese dessas discussões, os participantes pareciam ter estabelecido uma espécie de grupo permanente ou comunidade científica, centrada nos *campi* do MIT, de Stanford e de Carnegie-Mellon, diz Gardner (1996). Proeminente crítico desta abordagem, Varela (1991, p. 29, grifo do autor) afirma que

> A principal ideia que viria a impor-se no decorrer desta conferência foi o facto de a inteligência (inclusive a inteligência humana) se aproximar de tal forma daquilo que, intrinsecamente, é um computador e que a cognição pode ser *definida* pela *computação* de representações simbólicas.

A década de 1960 foi caracterizada pelo afloramento das sementes plantadas nos anos 1950 em discussões como as realizadas nas Conferências Macy. Fontes governamentais e privadas forneceram apoio financeiro significativo para o desenvolvimento da ciência cognitiva. Duas das figuras

principais para a consolidação dessa nova área, Jerome Bruner e George Miller, fundaram, em Harvard, o Centro de Estudos Cognitivos. Embora os projetos e produtos reais desse centro provavelmente não tenham sido indispensáveis para a vida desta área de pesquisa, durante aquele período praticamente não havia nenhuma pessoa jovem trabalhando nesse campo que não tivesse sido influenciada pela presença desse centro, pelas ideias que eram debatidas por lá e pela forma como elas eram implementadas em pesquisas subsequentes.

De acordo com Gardner (1996), George Miller, seu colega neurocientista Karl Pribram e Eugene Galanter abriram a década com um livro de grande impacto na psicologia e áreas relacionadas, intitulado *Plans and the structure of behavior*. Nele, os autores apresentam um enfoque cibernético do comportamento, em termos de ações, *feedback* e reajustes das ações conforme a retroalimentação. Em linhas gerais, esses três cientistas propunham legitimar, na prática, o abandono da discussão de estímulos e respostas, em favor de modelos mais abertos, capazes de interações propositais. No decorrer da mesma década, surgiam inúmeros "exemplares" decorrentes dos últimos desenvolvimentos. Estes recentes livros e outras publicações serviram como base para a formação dos novos cientistas cognitivos que, por sua vez, consolidavam a nova comunidade científica.

Segundo Gardner (1996), no final de 1969, já era possível pensar em uma abordagem de ciência cognitiva como um todo. Quando o nível de atividade em um campo chega a esse ponto, com uma comunidade aparentemente entusiasmada em torno dos avanços iminentes, geralmente se tem a consolidação de algum tipo de organização.

Em 1970, estava ocorrendo uma série de eventos que, para a ciência cognitiva, ocasionaram grandes avanços, através de uma fundação privada sediada em Nova York – a Fundação Alfred P. Sloan. Ela financiava o que ela mesma denominava "Programas Particulares", nos quais investia um montante considerável de dinheiro, em uma área, por um período de alguns anos, esperando com isso estimular algum progresso significativo no desenvolvimento da área de pesquisa.

No início de 1970, um Programa Particular havia sido implementado nas neurociências. Após esse evento, a Fundação Sloan estava à procura de um campo análogo, de preferência dentro das ciências

naturais, no qual pudesse investir uma quantidade semelhante de recursos. Em 1975, a fundação estava estudando o apoio a programas de vários campos e o Programa Particular na incipiente ciência cognitiva era o principal deles. No ano seguinte, foram promovidas algumas reuniões, nas quais os principais nomes que viriam a ser denominados cientistas cognitivos, oriundos de diferentes ciências, que também viriam a ser chamadas cognitivas, expuseram suas ideias. Embora a ciência cognitiva não estivesse tão madura quanto os neurocientistas da época acreditavam, apontavam os dirigentes da Fundação Sloan:

> [...] ainda assim, há muitas indicações, confirmadas pelas autoridades envolvidas nas investigações iniciais, de que muitas áreas das ciências cognitivas estão convergindo, e, além disto, há uma necessidade igualmente importante de desenvolver linhas de comunicação de uma área a outra, a fim de que instrumentos e técnicas de pesquisa possam ser compartilhados na construção de um corpo de conhecimento teórico [...] (GARDNER, 1996, p. 50).

Existe uma semelhança entre os estímulos fornecidos pela Fundação Macy à geração da cibernética e a iniciativa que a Fundação Sloan teve, com a ciência cognitiva. Depois de deliberado, a Fundação Sloan decidiu dar início a um programa de cinco a sete anos, envolvendo o investimento de até 15 milhões de dólares, o qual acabou sendo elevado para 20 milhões de dólares. Com base nesses financiamentos, dentro de pouco tempo o periódico *Cognitive Science* foi fundado, tendo seu primeiro número publicado em janeiro de 1977; logo em seguida, em 1979, uma sociedade de mesmo nome foi criada. A sociedade promoveu o seu primeiro encontro anual, em La Jolla, Califórnia, em agosto de 1979.

Com o passar do tempo, programas, cursos e boletins informativos estavam espalhados por todo o mundo. No entanto, havia discordâncias sobre o que tratava o campo da ciência cognitiva, quem a atendia, quem a ameaçava de ruir e em que direção ela deveria seguir. Com base nessas controvérsias, a fundação solicitou um relatório, em 1978, com o objetivo de explicitar tais questões. Esse Relatório do Estado de Arte [*State of the Art Report*] foi redigido pelos principais estudiosos do campo, com a colaboração de vários conselheiros. Os autores elaboraram uma figura com

as inter-relações entre os seis campos constituintes, chamada de hexágono cognitivo, que especificava o rol de pelo menos parte das ciências cognitivas:
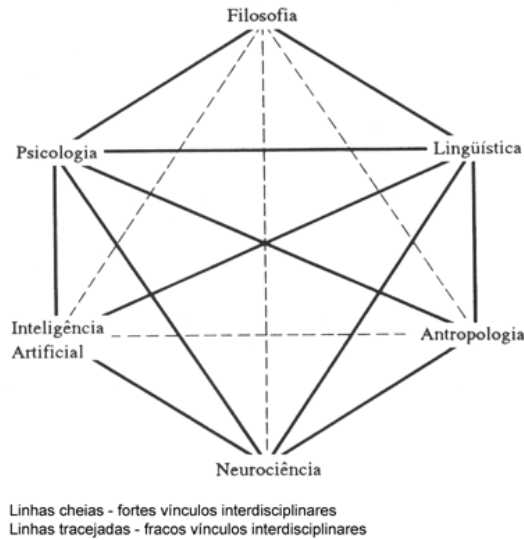
Figura 1. Hexágono cognitivo
Fonte: GARDNER, 1996, p. 52.

Foi realizado um esforço para indicar as conexões entre os campos que atuavam na ciência cognitiva e as possíveis conexões entre eles, expressos, na figura acima, nas linhas cheias e pontilhadas. A imagem e o relatório pretendiam disponibilizar um exame das principais linhas de pesquisa da ciência cognitiva. No entanto, a comunidade, de modo geral, teve uma recepção extremamente negativa do relatório. Segundo Gardner (1996), a reação negativa foi resultado do fato de que cada leitor teve uma interpretação do documento sob o prisma de suas próprias disciplinas. Naquele período, aparentemente, dada a inexistência de um paradigma de pesquisa estabelecido, os cientistas cognitivos tendiam às suas próprias inspirações, baseados em suas próprias convicções, formação acadêmica, intuições oriundas de suas experiências e suas práticas em suas áreas de pesquisa. Em virtude desses fatores, provavelmente, não era possível, em 1978, escrever um documento que obtivesse total apoio, ou mesmo uma concordância substancial dos principais integrantes do grupo

de pesquisadores, o que significaria a constituição de uma comunidade científica.

Em uma primeira análise, parece que, no princípio da história dessa área, muitas disciplinas compunham a base teórica daquilo que até o momento era chamado de ciência cognitiva. Esse é mais um elemento a corroborar a nossa hipótese exposta no final do capítulo anterior de que aquilo que chamamos de ciência cognitiva é decorrente do movimento da cibernética e que a ciência cognitiva não dispõe de um paradigma bem estabelecido em seus momentos iniciais, o que a faz começar como a grande maioria das ciências, por um estágio, na melhor das hipóteses, de pré-ciência. Nestes primeiros momentos, a ciência cognitiva dispunha de uma dinâmica própria que visava, de alguma forma, as aspirações de cada pesquisador ou grupo de pesquisadores. Havia certa discordância entre os primeiros membros oriunda da própria metodologia presente desde o período das Conferências Macy.

Partimos do princípio de que as disciplinas que constituem o momento de transição entre a cibernética e a ciência cognitiva não refletem necessariamente um modelo bem fixado de uma comunidade científica e, por sua vez, esse primeiro momento revela o ambiente das ciências cognitivas, que era um conjunto de disciplinas unidas, com suas especialidades e com o objetivo de investigar as questões atreladas à cognição.

Boa parte dos cientistas cognitivos visava o progresso da comunidade científica a partir das especificidades de suas áreas de pesquisas. Assim, por exemplo, os neurocientistas estavam muito mais preocupados com as estruturas dos neurônios e com os processos e relações neuronais do que com os aspectos fundamentalmente abstratos, subjacentes ao exame da cognição. O desenvolvimento disforme da pretensa comunidade científica suscitou a produção de exemplares que careciam de coerência com as demais estruturas desenvolvidas simultaneamente. Por sua vez, havia exemplares que diziam tratar da ciência cognitiva, mas não abarcavam as elucidações propostas pela antropologia, os estudos desenvolvidos pela linguística ou mesmo compartilhavam dos mesmos pressupostos, sejam metafísicos, sejam filosóficos. Nem todos, por exemplo, concordavam com a perspectiva funcionalista da mente como abordagem subjacente ao

tratamento dos processos cognitivos, que envolvia aspectos internalistas e representacionistas, conforme tratamos a seguir

3.2 Modelos e representações na ciência cognitiva

Como já enunciado no segundo capítulo, uma das características metodológicas centrais da ciência cognitiva na sua fase inicial era o uso de modelos na tentativa de oferecer abordagens explicativas de processos cognitivos. Ademais, boa parte, senão a totalidade dos pesquisadores, defendia a ideia de que os processos cognitivos pressupõem ou envolvem representações mentais.

A concepção e o uso de modelos aparecem com certa constância na atividade científica. Muitos campos, como, por exemplo, a física, a biologia e a astronomia, representam suas entidades ou fenômenos de investigação a partir de esquemas capazes de instaurar uma relação de equivalência ou de isomorfismo com os seus objetos de estudo.

Conforme assinala Dupuy (1996, p. 23): “O modelo científico é uma imitação humana da natureza que o cientista logo toma como ‘modelo’ – no sentido comum – desta.” Os modelos permitem, até certo ponto, com o uso de ferramentas matemáticas, um controle explicativo e preditivo capaz de sugerir novas experiências e formular hipóteses inéditas sobre um dado problema, um fenômeno, uma entidade. Continua Dupuy (1996, p. 27): “Conhecer é produzir um modelo do fenômeno e efetuar sobre ele manipulações ordenadas. Todo conhecimento é reprodução, representação, repetição e simulação.”

Os modelos, apesar de imitar o modelado, com frequência apresentam uma performance própria com uma dinâmica “desligada” e baseada na realidade. Em geral, eles costumam ser mais controláveis e fáceis de se manipular do que os fenômenos do mundo real. Há, por isso, um cuidado para evitar que o modelo se torne o objeto exclusivo de análise dos cientistas, podendo se desconectar da relação epistemológica com o fenômeno do mundo investigado. Talvez seja este exatamente um dos pontos da atual ciência cognitiva que, aos poucos, de modo consciente ou não, foi alterando seus objetivos centrais. Atualmente, parece-nos que essa área de pesquisa perdeu um pouco o objetivo epistemológico referente à explicação de processos cognitivos para a criação de autômatos

capazes de reproduzir, simular e emular processos cognitivos. O resultado são os androides ou robôs humanoides construídos por empresas de tecnologias, em poucos anos indistinguíveis do ser humano, pelos menos em sua aparência, física e psicológica. Mas essa é outra história, para outro momento.

Embora os modelos sejam utilizados com frequência na atividade científica, é necessário estar atento no tocante a saber quando efetivamente eles oferecem abordagens explicativas de um fenômeno. Segundo Pessoa Junior (2016, p. 102):

> Pode-se dizer que um modelo matemático ou computacional "captura" a realidade? Tomemos como exemplo a simulação computacional de um furacão. O que está sendo capturado são as relações entre as partes da atmosfera, ou seja, a estrutura dinâmica da realidade. Com a informação armazenada no computador, podem-se prever aproximadamente os efeitos causais de situações reais e também de situações contrafactuais, como por exemplo, o que aconteceria com um avião que entrasse em um furacão.
>
> [...]
>
> O modelo captura apenas a organização ou estrutura do sistema, e não os elementos em si, não a materialidade do sistema representado. Todo modelo tem sua materialidade própria, sejam os dispositivos e circuitos elétricos de um computador, sejam as estruturas celulares em um encéfalo (segundo a concepção materialista do problema mente-corpo). Mas se a materialidade do sistema que representa for distinta da do objeto que é representado, esta distinção estabelece um limite "qualitativo" para a modelagem.

Os modelos também podem apresentar questões ou dificuldades associadas ao significado, os quais dizem respeito à forma como poderiam surgir e em que consistiriam os conteúdos ou significados de sistemas representacionais. Enfim, como tais sistemas poderiam referir-se às coisas do mundo, ou melhor, como sistemas poderiam exibir intencionalidade intrínseca (SEARLE, 1980).

Os dois modelos mais utilizados na ciência cognitiva para simular e explicar os processos cognitivos são as máquinas do tipo Turing e as Redes Neurais Artificiais. Trataremos de ambos mais tarde neste capítulo,

ao falarmos do cognitivismo e conexionismo. Antes disso, enunciamos rapidamente um elemento fundamental no tratamento dos processos cognitivos neste momento: as representações.

Na ciência cognitiva, a noção de representação está relacionada ao binômio e disputa entre o internalismo e o externalismo. De modo geral, segundo o internalismo, a mente ou, para evitar debates e comprometimentos ontológicos no momento, os processos cognitivos estariam, de alguma forma, dentro do indivíduo, enquanto, no externalismo, tais processos se encontrariam, de algum modo, fora dele. Para os internalistas mais radicais, tais como os defensores de abordagens dualistas como a cartesiana, a mente existe independente do mundo externo e ela pode ser compreendida sem qualquer necessidade de recorrermos ao mundo externo.

Na visão dos internalistas radicais, elementos como dor, sentimentos e crenças seriam, em menor ou maior grau, independentes do corpo. Nesse sentido, a mente seria capaz de instanciar estados como uma dor no braço sem necessariamente ter um braço, tal como ocorre em alguns casos documentados pela medicina de "membros fantasmas".

Putnam (1975), embora não seja um externalista radical, fornece um exemplo contrário aos internalistas radicais, pretendendo mostrar, com um exemplo fictício, chamado "Argumento da Terra Gêmea", que os estados mentais não determinam, por exemplo, a referência ou os conteúdos do pensamento. Com este experimento mental, o seu autor procura mostrar que o conteúdo é externo à mente. Na Terra Gêmea, existe uma cópia idêntica a todos os indivíduos de nosso planeta. Nossos clones agem, falam, pensam, acreditam e fazem tudo da mesma maneira que nós. Todas as nossas propriedades internas são idênticas. Nesse exemplo, a palavra "água" é usada para fazer referência à substância idêntica, em todos os sentidos, com aquela que possuímos na Terra original. Entretanto, existe uma particularidade da Terra Gêmea, no sentido de que o líquido chamado "água" não é descrito pela fórmula H^2O, mas por um aglomerado de elementos que o autor descreve como XYZ.

Putnam (1975) procura mostrar, com esse exemplo, que os conceitos não são suficientes para determinar a referência do conteúdo e, embora nossos estados mentais sejam iguais, nós nos referimos a coisas distintas. Podemos notar que, quando eu e minha cópia falamos sobre "água", cada um de nós faz referência a um composto químico diferente.

Na concepção de Frawley (2000), que busca oferecer uma concepção intermediária entre os extremos nesse caso, a divisão feita entre internalismo e externalismo serve apenas para organizar a preferência explanatória de cada pesquisador. Frawley (2000, p. 15, grifo do autor) diz:

> [...] o internalismo e externalismo se referem às ideologias científicas que colocam o maior peso da explicação em unidades seja interna, seja externa ao pensamento. O *internalismo* não pode ser igualado ao nativismo, que é apenas um tipo de explicação internalista, e, também não é o mesmo que computacionalismo, embora tais explicações da mente geralmente enfatizem os fatos dentro do limite mente-mundo. O *externalismo* privilegia o externo em relação ao interno e ocorre em muitas formas – empiricismo, behaviorismo ou alguma escola de pensamento dessa natureza que considera fatores externos à mente. Mas é possível ser um externalista *e* um computacionalista, como atesta qualquer uma das muitas novas teorias da mente contextualizada.

Boa parte dos internalistas acredita na existência de um elemento mediador, chamado representação mental, entre o sujeito e o objeto de conhecimento. Para outros, a relação com o mundo ocorre de maneira direta. Essas duas vertentes são denominadas, respectivamente, representacionistas e não representacionistas.

Para os representacionistas, a representação funciona como uma ponte entre a mente e o mundo. "Uma representação é uma versão modificada do mundo", define Charniak (1993, p. 8, tradução nossa). O que caracteriza uma representação é sua propriedade de ser algo que pode se colocar no lugar de "outra coisa" (como um mapa, por exemplo). A principal característica definidora da ciência cognitiva seria, por conseguinte, a ideia de simuladores e/ou reprodutores de propriedades mentais, e, para que isso se efetive, as noções de representação e computação são fundamentais, durante o processo de modelagem computacional, como salienta Fodor (1980, p. 31, tradução nossa): "Sem representação, não há computação; sem computação, não há modelagem." Nesse sentido, mais do que uma escolha metafísica, a defesa de um internalismo representacionista também significava, de alguma forma, uma necessidade metodológica para o

tratamento de processos cognitivos para uma ala, muito significativa, da ciência cognitiva em sua fase inicial.

No contexto da ciência cognitiva, aos sistemas a que se atribuem representações (os quais podem ser artefatos tanto do cognitivismo quanto do conexionismo), pode-se enfatizar que as representações seriam dotadas de um conteúdo cuja natureza explicativa e ontológica varia dependendo da abordagem. Mais do que isso, a representação, sobretudo, guia o comportamento ou a atividade de um sistema. Haselager (2004, 106) assevera que os adeptos do representacionismo entendem que

> As duas características mais importantes das representações são que elas se colocam no lugar de algo e que o sistema usa as representações com o objetivo de guiar seu comportamento. De acordo com a ciência cognitiva tradicional, então, as representações desempenham um duplo papel: carregam um conteúdo e causam o comportamento. Mesmo se a ciência cognitiva clássica e o conexionismo discordam a respeito do formato das representações, eles têm esse pressuposto em comum.

O uso das representações implica um compromisso coletivo da comunidade científica da ciência cognitiva. Do ponto de vista da matriz disciplinar, o uso das representações é compartilhado pela comunidade de pesquisadores, todavia, a sua aplicação pode ser relativa a cada perspectiva presente no paradigma geral.

As representações são concebidas como um modo abstrato de reter conhecimentos sobre o mundo, por meio de símbolos, esquemas, imagens, ideias. A representação mental é uma espécie de imagem mental de algum objeto ou fenômeno do mundo. Os representacionistas afirmam que podemos reconhecer os objetos do mundo porque deles possuímos algumas representações em nossa mente.

A natureza das representações também é cercada de problemas. Seriam estas baseadas em símbolos bem estruturados ligados a uma rígida articulação sintática e semântica (cognitivismo)? Ou seriam elas representações distribuídas, fundadas em pesos ajustados mediante treinamento (conexionismo)? Seriam ambas? De mais algum tipo além dessas? Sem pretendermos dar conta desse problema da ciência

cognitiva, podemos sublinhar, segundo Thagard (1998), apenas que as estruturas representacionais da Inteligência Artificial, bem como das redes conexionistas, são complementares, em vez de competitivas.

A operacionalidade, por sua vez, sofre da dificuldade em operacionalização do próprio conceito de representação na ciência cognitiva. Ou seja, em qualquer ciência os conceitos devem ser aplicáveis, e em ciência cognitiva, não parece ser claro a quais sistemas deveria se aplicar a noção de representação.

Além da circunstância de que o próprio observador pode gerar problemas, esse problema é ilustrado, segundo Haselager (2004), pelo seguinte exemplo: ao observar a complexidade do caminho de uma formiga na areia da praia, um cientista cognitivo poderia ser tentado a considerar a complexidade da trilha como efeito de representações mentais "na cabeça" da formiga. Imagina-se que o cientista observa representações, quando possivelmente isso é desnecessário.

A relação entre mediação e local de ação da mente cria uma série de possibilidades. Supondo-se a possibilidade de um paradigma da ciência cognitiva, existe uma tendência de que seus adeptos são representacionistas, ou seja, eles acreditam na existência de algo mediador entre o programa e seu ambiente externo. Para eles, a representação consiste em um conjunto de símbolos adquiridos, a partir de alguma mediação, como sensores, os quais são capazes de expressar algo que está acontecendo com eles mesmos ou com o ambiente externo.

Procuramos mostrar nesta seção que os primeiros cientistas cognitivos em geral concordavam com o pressuposto metodológico do uso de modelos para o estudo de processos cognitivos. Também aceitam, de modo semelhante, que a representação exerce um papel essencial tanto do ponto de vista metodológico quanto do ponto de vista epistemológico e até ontológico desses fenômenos. No entanto, as semelhanças, se é que existem realmente, terminam por aqui. As estranhezas podem ser explicitadas na descrição de duas das mais fortes vertentes da ciência cognitiva em sua fase inicial: o cognitivismo e o conexionismo, dos quais tratamos a seguir.

3.3 Cognitivismo

No capítulo anterior mostramos a existência de duas frentes ideológicas na constituição da ciência cognitiva, oriundas de duas obras capitais publicadas 1943: *Behavior, Purpose and Teleology*, de Norbert Wiener, Arturo Rosenblueth e Julian Bigelow, e *A logical calculus of the ideas immanent in nervous activity*, de Warren McCulloch e Walter Pitts. Essas duas obras são o gérmen do que posteriormente seriam as duas principais vertentes da ciência cognitiva em sua fase inicial: o cognitivismo e o conexionismo, dos quais tratamos no restante deste capítulo. Seguindo a abordagem de Kuhn, avaliamos em que medida a existência desses dois programas de pesquisa indicaria que a ciência cognitiva nesse estágio se configuraria como pré-ciência ou se, de alguma forma, poderia ser considerada ciência normal, avaliando se algum desses projetos, uma vez podendo ser considerados paradigmas, estivesse em evidência, ou seja, fosse dominante.

No decorrer de 1956 ocorreu o afloramento da primeira grande corrente da ciência cognitiva: o cognitivismo. Também denominado Inteligência Artificial, funcionalismo lógico-computacional, o cognitivismo, durante os primeiros momentos da ciência cognitiva, contou com o apoio de instituições de prestígio, revistas científicas e com a aplicação do programa em tecnologias.

Segundo Gardner (1996, p. 159),

> [...] o nome Inteligência Artificial foi pronunciado em 1956, quando alguns cientistas como John McCarthy, Marvin Minsky, Allen Newell e Herbert Simon discutiram as possibilidades de se produzir programas computacionais que pudessem "se comportar" ou "pensar" de maneia inteligente, como por exemplo solucionar problemas, reconhecer padrões, tomar parte em jogos e raciocinar logicamente. Estes cientistas baseavam-se na hipótese de que todo aspecto de aprendizagem ou de qualquer outra faceta da inteligência pode ser descrito de forma tão precisa que se pode fazer com que uma máquina o simule.

Para Charniak e McDermott (1985, p. 113, tradução nossa), "IA é o estudo de faculdades mentais por meio do uso de modelos

computacionais. Seu objeto é a mente, entendida como um sistema processador de informação". Já para Dreyfus (1993, p. 39, tradução nossa), um dos críticos dessa teoria, "A IA é a tentativa de simular o comportamento humano inteligente utilizando-se técnicas de programação que precisam demonstrar pouca ou nenhuma semelhança com os processos mentais humanos".

Varela, Thompson e Rosch (1991), outros críticos dessa corrente, assinalam que, nesse momento, o cognitivismo podia ser considerado como o centro ou núcleo da ciência cognitiva. Na visão de Varela, Thompson e Rosch (1991, p. 30, grifo do autor), a ferramenta mais importante do cognitivismo é a metáfora do computador digital:

> [...] cognitivismo consiste na hipótese de que a cognição – incluindo a cognição humana – é a manipulação de símbolos nos moldes daquilo que é executado pelos computadores digitais. Por outras palavras, a cognição é uma *representação mental*: a mente é definida como operando em termos de manipulação de símbolos que representam características do mundo ou representam o mundo como sendo de um determinado modo.

A grande motivação dos primeiros cientistas desta corrente da ciência cognitiva foi o fato de que, pela primeira vez, sistemas artificiais conseguiam realizar com sucesso comportamentos que até então eram de exclusividade humana. Dentre eles estavam a atividade de resolução de problemas do tipo lógico-matemáticos e a participação em jogos como xadrez.

A Inteligência Artificial pode ser historicamente dividida em duas versões: IA forte e fraca. De acordo com a primeira, a mente é um programa de computador. Ambos, computador e mente, devem ser concebidos como um sistema simbólico – entidade que processa, transforma, elabora e manipula símbolos de vários tipos, processando informações no decorrer do tempo, numa ordem mais ou menos lógica. Para essa versão, a mente está para o cérebro assim como o *software* está para o *hardware* do computador. Como afirmam Newell e Simon (1972, p. 19, tradução nossa), "[...] o homem é um sistema de processamento de informação, pelo menos quando está resolvendo problemas", e um computador pode perfeitamente simular tal sistema.

A IA em sua versão fraca não adota a hipótese da identidade entre mente e programa. Seus adeptos afirmam existir apenas uma semelhança entre mente humana e programa computacional. Os programas de computador são um bom modelo da mente, podendo explicar seu funcionamento e suas características. Tanto a mente quanto o programa manipulam símbolos e seguem regras lógicas. Porém, não são exatamente o mesmo objeto.

A noção do funcionamento mental comparado ao funcionamento computacional forneceu um meio muito poderoso para se abordar a cognição. O argumento cognitivista concebe que o comportamento inteligente pressupõe a capacidade de representar o mundo de alguma forma. Dessa maneira, o comportamento cognitivo só pode ser explicitado partindo-se do princípio de que os sistemas atuam representando as mais diversas situações do mundo, conforme já discutido no capítulo anterior.

Os cognitivistas não defendem que, se tivéssemos de abrir a cabeça de alguém e olhar o cérebro, encontraríamos símbolos sendo manipulados. Embora seja fisicamente realizado, o nível simbólico não é redutível ao nível físico. Desse modo, os mesmos símbolos podem ser implementados em numerosas formas físicas capazes de instanciá-los. O cognitivismo defende um nível simbólico irredutível e distinto de sistemas físicos na explicação da cognição. Uma vez que os símbolos são elementos semânticos, os cognitivistas supõem um terceiro nível que assinala o semântico e o representacional. Desse modo, um dos problemas principais do cognitivismo está na correlação entre os estados intencionais, como as crenças e os desejos, com as mudanças físicas de um agente. Em linhas gerais, se os estados intencionais possuem propriedades causais sobre os estados físicos, é importante mostrar em que medida esses estados são capazes de determinar o comportamento de um sistema. A noção de computação simbólica, na perspectiva cognitivista, pressupõe que os símbolos têm uma realidade simultaneamente física e semântica.

Essa vertente tende a ser representacionista, visto que seus adeptos pressupõem a existência de algum elemento mediador entre o programa e o ambiente externo. Seja por intermédio de sensores, seja por outros mecanismos, quando a máquina recebe algum símbolo, seja por meio do ambiente, seja por si mesma, é capaz de o reconhecer e fazer referência à sua representação.

Os modelos do cognitivismo para a explicação dos processos cognitivos são as máquinas do tipo Turing. Inicialmente, Turing estava voltado a resolver problemas de computabilidade, mais especificamente, de definir funções computáveis. Para resolver essa e outras questões em torno delas, propôs, em 1936, uma definição de algoritmo. Posteriormente, tal sistema formal foi utilizado tanto como modelo para o estudo de processos cognitivos quanto serviu de base para a construção dos atuais computadores digitais.

As máquinas de Turing são geralmente distinguidas entre finitas e infinitas. Ambas possuem as mesmas especificações técnicas, diferenciando-se basicamente apenas entre suas capacidades de armazenamento de dados. As máquinas finitas são aquelas que apresentam memória limitada, enquanto as infinitas possuem memória potencialmente infinita, sendo capazes de computar funções mais complexas do que as outras, tais como operações aritméticas como a soma de dois números naturais quaisquer.

Turing (1936, p. 231, tradução nossa) assim apresenta a máquina:

> Podemos comparar um homem no processo de computação de um número real [uma função] com uma máquina que é apenas capaz de um número finito de condições q1, q2, ..., qR chamadas "m–configurações" [estados internos]. À máquina pertence uma "fita" (análoga a um papel) que a percorre, e é dividida em seções (chamadas quadrados), cada uma capaz de armazenar um "símbolo". Em cada momento há apenas um quadrado, o r–ésimo, guardando o símbolo Ø(r) que está "na máquina", chamado "quadrado lido". O símbolo sobre ele é denominado o "símbolo lido", o único com o qual a máquina está, por assim dizer, "diretamente ligada". Contudo, ao alterar uma m–configuração, a máquina pode efetivamente relembrar algum dos símbolos que ela "viu" (escaneou) anteriormente. Seu comportamento possível em qualquer momento é determinado pela m–configuração qn e o símbolo lido 124 Ø(r). Este par será chamado de "configuração", que determina o possível comportamento da máquina. Em algumas das configurações nas quais o quadrado lido está vazio (não carrega nenhum símbolo), a máquina escreve sobre ele um novo símbolo: em outras configurações ela apaga o símbolo lido. Pode ainda mudar o quadrado que está sendo lido, mas somente deslocando-se um lugar à direita ou à esquerda. Em adição a qualquer uma destas operações, a m–configuração pode

> ser mudada [...] Se para cada estágio o movimento da máquina (no sentido acima explicitado) é completamente determinado pela configuração, podemos chamá-la de uma máquina automática [...] Se uma máquina automática imprime dois tipos de símbolos, dos quais o primeiro tipo (chamado figuras) consiste inteiramente de 0 e 1, (os outros sendo chamados símbolos do segundo tipo), então ela é uma máquina computadora.

A máquina de Turing é constituída basicamente por um conjunto finito *S* de símbolos, um conjunto finito *Q* de estados internos, uma fita de memória potencialmente infinita, um conjunto *P* finito de instruções e um agente que realiza as instruções. Dessa maneira, o funcionamento da máquina é determinado conforme as instruções compostas no seu interior, visando realizar alguma função computável. A estrutura das instruções é a seguinte: “”. “” representa o estado atual da máquina, “” representa o símbolo lido da máquina “” representa o novo símbolo da máquina, “D”, “E” e “I” representam a direção da cabeça de leitura da máquina (direita, esquerda ou imóvel), “” representa o novo estado da máquina.

Como descreve Turing (1950), para cada estado da máquina existem instruções fundamentais para iniciar a computação das funções. Dessa maneira, é fundamental que a memória da máquina seja potencialmente infinita. A figura abaixo ilustra a estrutura da máquina de Turing.
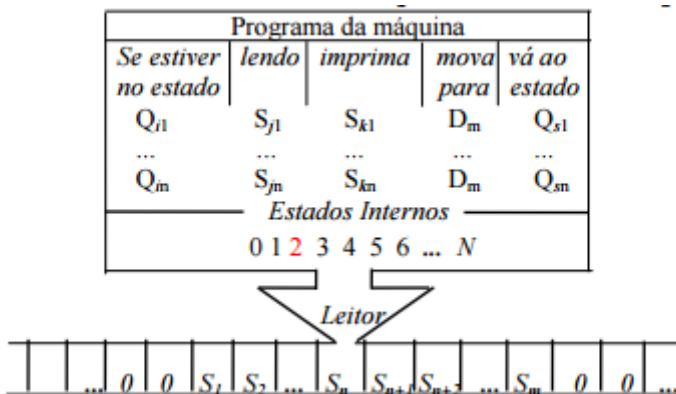
Figura 2. Estrutura geral da máquina de Turing
Fonte: ALVES, 1999, p. 61.

A máquina possui uma fita dividida em quadrados, cada um capaz de armazenar um símbolo. Em cada momento a máquina manipula apenas um único quadrado, podendo apagá-lo, inserir um novo símbolo ou deixá-lo inalterado. Conforme as suas configurações, a máquina pode se "recordar" de um símbolo anteriormente lido. O seu movimento é determinado pelas configurações expressas pelos estados internos e o símbolo lido. O par constituído do símbolo e do estado interno é nomeado configuração.

Abaixo, apresentamos um exemplo particular de uma máquina de Turing, denominada somadora unária. Ela soma dois números M e N na notação unária, ou seja, com base 1, no qual 0 = 0, 1 = 1, 2 = 11, 3 = 111 etc.

```
0-0-0-D-2    1-1-1-D-1    2-A-0-D-2
0-1-0-D-1   1-A-1-D-2
```
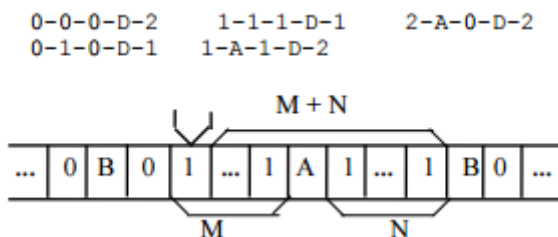
Figura 3. Soma entre M + N, conforme o modelo da máquina de Turing.
Fonte: ALVES, 1999, p. 64.

Conforme Alves (1999), a computação da soma inicia-se a partir do estado 0, lendo o primeiro símbolo à esquerda de M. Caso M seja 0, a máquina move o leitor para a direita e vai ao estado 2, como indicado pela primeira instrução da esquerda para a direita do programa acima. Após isso, substitui A por 0, move o leitor para a direita, como dito na quinta instrução. O leitor vai ficar no estado dois lendo 1 ou 0 e nenhuma das instruções de seu programa possui 0–0 ou 0–1 como configuração. A resposta, neste caso, será o próprio N. Caso M não seja 0, a máquina substitui o primeiro símbolo de M por 0 e vai à direita, indo ao estado 1, como manda a segunda instrução. Assim, o leitor busca a letra A, seguindo a terceira instrução. Substitui-se esta letra pelo número 1. Isso é o que manda a quarta instrução. A resposta será a sequência de números 1's que restou na fita.

Já com o objetivo de tratar da inteligência e não mais propriamente de questões de computabilidade, Turing (1950) formula as primeiras discussões envolvendo as aplicações da informação a uma teoria

adequada que pretende analisar a inteligência e o pensamento, segundo uma concepção mecanicista formal. A sua proposta consiste na defesa do pressuposto de que "[...] pensar é calcular" (TURING, 1950, p. 436, tradução nossa). Essa tese parte do pressuposto de que, se aceitarmos que a atividade de pensar é, preponderantemente, constituída pela manipulação eficiente de informação/dados, por meio de regras lógicas, poderíamos sustentar que computadores também seriam capazes de pensar, se fossem capazes de tal manipulação. A tese de que "pensar é calcular" está pautada em um famoso jogo proposto por ele, chamado "jogo de imitação", ou teste de Turing. O jogo, assevera Turing (1950, p. 433, tradução nossa):

> É jogado por três pessoas: um homem (A), uma mulher (B), e um interrogador (C), que pode ser de qualquer dos sexos. O interrogador permanece num quarto, separado dos outros dois. O objetivo do jogo, para o interrogador, é determinar, em relação aos outros dois, qual é o homem e qual é a mulher. Ele os conhece por rótulos X e Y e no fim do jogo dirá ou "X é A e Y é B", ou "X é B e Y é A".

O objetivo do homem é, justamente, tentar induzir o interrogador a fazer uma identificação errada. Já o objetivo da mulher é ajudar o interrogador a fazer a escolha certa. Sua melhor estratégia é possivelmente responder a todas as perguntas de maneira correta. Como ressalta Turing (1950, p. 433, tradução nossa), ela poderia dizer coisas como: "Eu sou a mulher, não o escute!". Mas isso seria inútil, porque o homem também poderia dar uma resposta semelhante.

O objetivo do jogo é avaliar em que medida, através de perguntas e respostas, ao substituir um dos interlocutores por uma máquina, ela seria capaz de enganar um interrogador. Ao final, se a máquina fosse capaz de responder às perguntas do interrogador e conseguisse confundi-lo, de sorte que ele julgue que as respostas foram dadas por um ser humano, então ela passaria no teste, ou seja, ela teria capacidade de pensar. Turing (1950) destaca que poderíamos substituir a pergunta "Podem as máquinas pensar?" por "Existem computadores digitais imagináveis com bom desempenho no jogo de imitação?". Segundo sua concepção, ambas as perguntas equivalem à ideia de que um computador digital, programado de forma adequada, pode ser preparado para desempenhar satisfatoriamente o papel do jogador A no jogo de imitação.

A máquina de Turing é um modelo abstrato mecânico capaz de manipular símbolos segundo regras específicas. Em cada instante há apenas um símbolo sendo manipulado. A máquina, a partir da sua configuração algorítmica, pode recuperar símbolos lidos anteriormente. Seus procedimentos possíveis são resultantes da relação entre os símbolos lidos e suas configurações. Ela computa funções, de modo que, fornecidos os elementos do domínio de uma função desejada, ela encontrará uma resposta adequada, se bem construída, de acordo com os padrões preestabelecidos no momento de sua configuração.

Para Putnam (1960), o desenvolvimento das noções da máquina de Turing e a invenção do computador ajudaram a resolver ou a dissolver o clássico problema da relação mente-corpo. O respectivo problema clássico faz referência ao fato de que os fenômenos mentais parecem ser qualitativa e substancialmente diferentes dos corpos que parecem instanciá-los.

Na concepção funcionalista, subjacente à proposta de boa parte dos cognitivistas, os diferentes programas, em computadores iguais ou diferentes, poderiam ser executados em diferentes estruturas adequadamente construídas para processar informações. Os funcionalistas entendem que a mente é um sistema processador de informações que manipula os símbolos de entrada (*inputs*) do sistema gerando uma saída (*output*), seguindo certas regras lógicas. Assim, o conjunto de operações lógicas (*software*) poderia ser descrito independentemente do *hardware* específico no qual ocasionalmente houvesse sido instanciado. A analogia dos computadores com o sistema humano sugere que o seu cérebro corresponde ao *hardware*, enquanto os padrões de processamento de informações, ou seja, o pensamento, fazem referência ao *software*. Além disso, os seres humanos, não menos que os computadores, armazenam programas, e as mesmas linguagens simbólicas podem ser invocadas para descrever programas de ambas as entidades.

Um dos objetivos principais da IA é construir modelos computacionais capazes de simular comportamentos que, se fossem realizados por seres humanos, seriam inequivocamente considerados inteligentes. Segundo Feigenbaum e Feldman (1968, p. 3, tradução nossa), o objetivo dos cientistas da IA é “[...] construir programas de computador que exibem comportamentos que são chamados de inteligentes quando observados em seres humanos”. Dentre estes cientistas encontramos Newell, Shaw e Simon (1958), Putnam (1967) e Minsky (1967).

Para alguns pesquisadores da IA, como os citados acima, a inteligência é uma questão de aprendizado, de aquisição de memória ou conhecimento-base de uma extensão suficiente, e de desenvolver os mecanismos de recuperação necessários para usá-lo (SCHANK; BIRNBAUM, 1997). Para ser mais inteligente, um sistema deve aumentar seu conhecimento, proporcionado pela introdução de programas mais complexos. Para esses cientistas, é possível a construção de entidades inteligentes fazendo com que elas realizem comportamentos inteligentes. Segundo Schank e Birnbaum (1997, p. 78),

> [...] nós e outros cientistas da IA sustentamos que se podem construir entidades inteligentes analisando em que consiste o comportamento inteligente, determinando as regras que governam esse comportamento e implementando tais regras em uma máquina.

A ideia desses cientistas é que comportamentos inteligentes podem ser explicados de um modo mecânico. Por isso, os programas da IA, acreditam seus criadores, além de simular comportamentos humanos inteligentes, explicam os processos pelos quais o homem passa na resolução de problemas.

Na IA, um comportamento inteligente é aquele cujo resultado envolve a capacidade de resolução de um problema do melhor modo possível. O comportamento inteligente, dizem Feigenbaum e Feldman (1968, p. 6, tradução nossa),

> [...] seja do homem, seja da máquina, será aquele que, dada determinada situação, precisa escolher a possibilidade correta para ela. Deste modo, para a máquina agir inteligentemente, precisa pesquisar as diversas incertezas do problema de um modo altamente seletivo, explorando caminhos relativamente férteis, com soluções, e ignorando caminhos relativamente estéreis.

Pesquisadores como Newell, Shaw e Simon (1958) procuraram simular processos cognitivos segundo os quais poderia ser mais certa a possibilidade de estudo. Esses processos eram os referentes ao pensamento lógico-matemático. Segundo os pesquisadores acima referidos, o pensamento, enquanto relacionado aos processos cognitivos, poderia ser

explicado mecanicamente através de programas computacionais. Críticos da IA como Penrose (1993), por exemplo, não concordam que o pensamento, em especial o matemático, possa ser simulado computacionalmente. Segundo Penrose (1993, p. 128),

> O pensamento matemático não pode ser descrito computacionalmente porque nele estão contidas crenças, intuição, compreensão, sutileza, talento artístico etc. o pensamento matemático não pode ser reduzido ao cálculo cego, à pura manipulação de símbolos.

No pensamento encontramos intuição, compreensão, bom senso, elementos estes que, provavelmente não podem ser formalizados, acredita Penrose. Uma máquina, por exemplo, apenas segue regras e manipula símbolos, sem compreender o que faz. O pensamento matemático, ao contrário, requer uma boa dose de entendimento, sutileza e mesmo talento artístico. Ainda que em algumas vezes ele possa ser reduzido a um cálculo cego, onde o computador tem muito mais vantagens do que o ser humano, o pensamento em geral não pode ser computacionalmente simulado, afirma Penrose (1993).

A crítica ao estudo do pensamento matemático estende-se aos outros tipos de pensamento, segundo Penrose (1993, p. 112), uma vez que

> [...] não há nada de essencial que separe o matemático de outros tipos de pensamento, de modo que a nossa demonstração de que o entendimento matemático é algo que não pode ser simulado em termos computacionais pode ser concebida também como uma demonstração de que o próprio entendimento – um dos mais essenciais componentes da inteligência genuína – é algo que se situa além de qualquer tipo de atividade puramente computacional.

Outra crítica bastante devastadora, pelo menos em termos epistemológicos, ao projeto cognitivista foi feita por John Searle. O argumento criado por este filósofo para refutar a ideia da possibilidade de atribuição de inteligência (e estados e faculdades mentais em geral) às máquinas do tipo Turing foi um dos mais populares da história da ciência cognitiva.

Atribuir inteligência a um sistema, segundo a maioria dos cientistas da IA, significa ele ser capaz de realizar certos comportamentos, realizados através da manipulação de símbolos e seguimento de regras lógicas (NEWELL; SHAW; SIMON, 1958; PUTNAM, 1967). Searle (1984, p. 37) ironiza:

> O colega de Simon, Alan Newell diz “já descobrimos” (notem que Newell diz “descobrimos”, não “supusemos” ou “consideramos a possibilidade”, mas descobrimos) que a inteligência é justamente uma questão de manipulação de símbolos físicos; não tem nenhuma ligação essencial com qualquer tipo de material ou umidade biológica ou física. Antes, qualquer sistema capaz de manipular símbolos físicos de modo correto é capaz de inteligência no mesmo sentido literal que a inteligência humana dos seres humanos. Simon e Newell sublinham, pela sua honra, que não existe nada de metafórico nestas pretensões; proferem-nas de um modo inteiramente literal... Marvin Minsky do MIT diz que a próxima geração de computadores será tão inteligente que “teremos muita sorte se eles permitirem manter-nos em casa como animais de estimação domésticos”.

O objetivo de Searle (1984) é mostrar que a concepção acima é equivocada. Para isso, baseia-se no fato de que o computador digital funciona apenas sintaticamente. Segundo tal filósofo, isso não é suficiente para esta máquina compreender o que faz. Porém, a compreensão é condição necessária para a inteligência. Além disso, características fundamentais da mente humana como consciência e intencionalidade são causadas pelo cérebro e são propriedades dele, um sistema biológico com certas características físicas que permitem a emergência de fenômenos mentais, de processos cognitivos. Sistemas puramente formais não são capazes de originar ou possuir consciência e intencionalidade. Uma máquina do tipo Turing, conclui Searle (1984, p. 40), “[...] não pode ser uma mente porque esta possui mais do que uma estrutura formal, possui um conteúdo”.

Searle (1984) recria o teste de Turing para mostrar que os computadores digitais não pensam e não são inteligentes. Substitui a máquina implicada no teste por um ser humano funcionando do mesmo modo que ela, ou seja, manipulando símbolos e seguindo regras lógicas. Se Searle conseguir provar que o comportamento deste ser humano não

é inteligente, tampouco o comportamento da máquina o será, mesmo passando no teste de Turing. Para Searle (1984), a questão fundamental não está no fato da máquina de Turing ser ou não capaz de responder às questões do interrogador daquele teste. O problema encontra-se na falta de características fundamentais desta máquina, tais como a consciência, para a caracterização do pensamento ou da inteligência.

O argumento do quarto chinês é assim exposto por Searle (1984, p. 40):

> Imaginemos que alguém está fechado num quarto e que neste quarto há vários cestos cheios de símbolos chineses. Imaginemos que alguém, como eu, não compreende uma palavra de chinês, mas que lhe é fornecido um livro de regras em inglês para manipular os símbolos chineses. As regras especificam as manipulações dos símbolos de um modo puramente formal em termos da sua sintaxe e não da sua semântica. Assim a regra poderá dizer: "tire do cesto número um um símbolo esticado e ponha-o junto de um símbolo encolhido do cesto número dois". Suponhamos agora que alguns outros símbolos chineses são introduzidos no quarto e que esse alguém recebe mais regras para passar símbolos chineses para o exterior do quarto. Suponhamos que, sem ele saber, os símbolos introduzidos no quarto se chamam "perguntas" feitas pelas pessoas que se encontram fora do quarto e que os símbolos mandados para fora do quarto se chamam "respostas às perguntas". Suponhamos, além disso, que os programadores são tão bons para escrever programas e que alguém é igualmente tão bom em manipular os símbolos que muito depressa as suas respostas são indistinguíveis das de um falante chinês nativo. Lá está ele fechado no quarto manipulando os símbolos chineses e passando para fora símbolos chineses em resposta aos símbolos chineses que são introduzidos.

Searle então afirma que o comportamento do indivíduo do quarto e o de um nativo chinês são praticamente indistinguíveis. Porém, o indivíduo do quarto não entende uma só palavra do chinês. Assim, o comportamento deste indivíduo não pode ser considerado inteligente porque lhe falta a semântica, característica fundamental para um comportamento deste tipo ser inteligente. Para Searle (1984, p. 45),

> [...] pensar é mais do que apenas uma questão de eu manipular símbolos sem significado; implica conteúdos semânticos significativos. Estes conteúdos semânticos são aquilo que nós indicamos por "significado".

Embora o argumento de Searle esteja direcionado para a compreensão de uma língua, de fato ele pode ser aplicado aos estados e faculdades mentais em geral. Mesmo que consiga simular uma dor, por exemplo, um computador de fato não tem essa dor. Isto porque, dentre outras coisas, não possui intencionalidade e consciência, que não podem ser apenas originadas pela manipulação de símbolos, afirma Searle (1998, p. 82).

Na concepção de Searle (1984), processos cognitivos são causados pelo cérebro, emergem da interação entre neurônios. Os cientistas da IA não levam em consideração as propriedades biológicas do cérebro ao construir suas máquinas, afirma Searle (1984). Ao contrário, elas apenas funcionam manipulando símbolos e seguindo regras lógicas. Por isso, não podem ser mentes. Sendo assim, jamais poderão efetivamente apresentar processos cognitivos e não podem ser considerados bons modelos para a sua explicação. A seguir apresentamos a vertente conexionista, alternativa ao cognitivismo, que procura considerar os aspectos físicos e biológicos de um sistema no tratamento explicativo dos processos cognitivos.

3.4 Conexionismo

O conexionismo é a outra grande vertente da ciência cognitiva em seus primeiros passos. É também denominado de Redes Neurais Artificiais, PDP (processamento distribuído em paralelo) ou funcionalismo neurocomputacional, por muitos de seus expoentes, além de McCulloch (1965), já exposto anteriormente. Alguns dos principais nomes ligados a essa área de estudos são Rosenblatt (1962), Hopfield (1982), Kohonen (1987), Rumelhart e McClelland (1986) e Caudill e Butler (1992).

Essa perspectiva funda sua abordagem epistemológica na ideia de que o cérebro é o sistema em que ocorrem ou são possíveis a emergência de processos cognitivos. O cérebro humano é formado por mais de uma centena de bilhão de elementos computadores chamados neurônios. Essa rede de neurônios é responsável por todos os fenômenos que chamamos

pensamento, emoção e cognição. Desse modo, para simular ou explicar processos cognitivos humanos, é preciso levar em consideração o estudo do cérebro, suas características, funcionamento, suas partes constituintes e as relações estabelecidas entre elas.

Conexionistas como os acima citados têm por fim criar sistemas inspirados no cérebro humano para simular e explicar, dentre outras coisas, comportamentos humanos inteligentes. Porém, pesquisadores como Caudill e Butler (1992, v. 1, p. 4) e McCulloch e Pitts (1943, p. 117) reconhecem que as redes neurais artificiais (RNAs) são apenas uma aproximação muito limitada do cérebro humano. Um simples exemplo disso é o fato dos conexionistas chamarem as partes constituintes das RNAs de nódulos (*neurodes*) e não neurônios (*neuron*).

A mente, para os conexionistas, não é simplesmente um sistema manipulador de símbolos e seguidor de regras lógicas, como afirmam os cientistas da IA. Em vez disso, ela é entendida como um conjunto de neurônios relacionados entre si, produzindo estados mentais, originando conhecimento, aprendizagem, comportamento inteligente.

A abordagem conexionista dos processos cognitivos, segundo seus defensores, proporciona um novo modo de pensar sobre percepção, memória, aprendizagem, pensamento e sobre os mecanismos computacionais básicos para o processamento inteligente de informações em geral.

Como já dissemos, as bases do conexionismo foram estabelecidas por McCulloch (1965). Em linhas gerais, a sua ideia era representar cada atividade mental por alguma proposição lógica. No caso da dor, por exemplo, a rede a simulará através da conexão entre seus nódulos. Estas conexões serão equivalentes a uma determinada proposição temporal do cálculo proposicional da lógica clássica, conforme ilustra Alves (1999).

Segundo McCulloch e Pitts (1943), grande parte das atividades mentais poderiam ser descritas em termos de conexões e estas em termos de proposições lógicas. Logo, por transitividade, tais atividades poderiam ser descritas por meio de proposições lógicas. A ideia que realmente liga McCulloch ao conexionismo é a de análise dos fenômenos mentais através de conexões neuronais.

Os conexionistas admitem que seus modelos operam mais eficientemente com a percepção e outros processos de nível inferior. Segundo Gardner (1996, p. 417),

> Mesmo aqueles que simpatizam com abordagens PDP admitem que elas operam mais eficientemente com a percepção e outros processos "de nível inferior" (subsimbólicos) do que com solução de problemas de grande escala, detecção de problemas, invenção e outros empreendimentos "simbolicamente carregados". Como Rumelhart e seus colegas colocam sucintamente, o que é difícil descrever na estrutura PDP são "o processo do pensamento, os conteúdos da consciência, o papel dos processos seriais, a natureza dos modelos mentais, as razões para as simulações mentais e o importante papel sinergético da linguagem no pensar e na formação de nosso pensamento".

Nos modelos conexionistas, as informações são codificadas não em estruturas simbólicas, mas através dos padrões de ativação das conexões entre as unidades. Smolensky (1987) usa o termo subsimbólico para designar processos (estados) físicos que, de alguma forma, participam como substratos dos estados simbólicos abstratos. Ou seja, os subsímbolos são constituintes básicos dos processos simbólicos. São também menos primitivos na escala de abstração do sistema cognitivo. De outra forma, constituem uma estrutura intermediária entre os planos neural e simbólico.

Os conexionistas propuseram essa nova ideia sobre a representação e a computação, inspirados na analogia da estimulação dos neurônios e da ativação difusa. Ao passo que a proposta cognitivista tem como base um processamento de informação em série, a proposta conexionista trabalha com um processamento de informação em paralelo e distribuído, possibilitando fazer mais de uma operação ao mesmo tempo.

A perspectiva conexionista consiste na ideia de que a cognição resulta do processamento coletivo feito pelas várias unidades simples que compõem uma rede neural. Destacam Rumelhart e McClelland (1986, p. 10, tradução nossa):

> Esses modelos [conexionistas] assumem que o processamento de informação ocorre pela interação de um grande número de elementos processadores simples chamados de unidades, cada um enviando sinais excitatórios e inibitórios para os outros.

Como os cognitivistas, grande parte dos conexionistas tendem a ser representacionistas e internalistas. Diferentemente dos cognitivistas, porém, os conexionistas não entendem a representação como um conjunto de símbolos, mas como um padrão de conectividade, ou seja, a partir da relação entre as partes básicas da rede. Assinala Gardner (1996, p. 414):

> Em vez de operações seriais ou computações sobre símbolos ou cadeias de símbolos, em vez de "executivos", "intérpretes" e "unidades centrais de controle", a abordagem PDP [processamento distribuído em paralelo, ou conexionismo] tipicamente postula milhares de conexões entre centenas de unidades (em princípio, a abordagem pode ser estendida a milhões ou mesmo bilhões de conexões). As redes resultantes apresentam a sinalização de excitações e inibições de uma unidade para outra. "Percepção", "ação" ou "pensamento" ocorrem em consequência da alteração das forças (ou pesos) das conexões entre estas unidades. Uma tarefa é concluída ou um *input* processado quando o sistema finalmente se "acomoda" ou "relaxa" (pelo menos provisoriamente) em um conjunto satisfatório de valores ou "estados estáveis" – em suma, em uma "solução".

Segundo O'Reilly e Munakata (2000), o conexionismo consiste na ideia de que, para explicar a cognição, não basta apenas reduzi-la a elementos mais simples, como neurônios ou interconexões decorrentes de parâmetros excitatórios e inibitórios. É preciso explicar como a combinação destes elementos é capaz de produzir os processos cognitivos. Tal qual as engrenagens que interagem no interior de uma máquina, para se compreender o seu funcionamento é importante especificar como elas interagem para produzir fenômenos mais gerais. Assim, para se explicar processamento cognitivo, por exemplo, existe uma necessidade de se entender a forma como ocorre a interação entre o grande número de "elementos simples", chamados de neurônios, com a sua imensa quantidade de conexões, chamadas de conexões sinápticas. As RNAs são marcadas como uma das primeiras vertentes a produzir modelos do sistema nervoso com grau de precisão suficiente para poder se observar o comportamento emergente dos neurônios trabalhando em paralelo.

Segundo Kovács (2006), a origem da teoria das RNAs está atrelada aos modelos matemáticos e aos modelos da engenharia, os quais tomam como base os neurônios biológicos. Ao longo da história de

pesquisa dessa base biológica, observou-se a existência de manifestação elétrica entre os chamados neurônios biológicos. Nas últimas décadas, em decorrência do trabalho de vários pesquisadores, passou-se a compreender os neurônios biológicos como elementos processadores fundamentais do sistema nervoso, compostos de um grande número de entradas, chamadas de conexões sinápticas.

Grosso modo, nos sistemas biológicos de neurônios, os sinais que chegam são pulsos elétricos, denominados impulsos nervosos, cujas sinapses correspondem a regiões eletroquímicas entre neurônios, por onde existe a troca de estímulos por meio de substâncias conhecidas como neurotransmissores. O resultado dessa transferência de estímulos, dependendo do tipo de neurotransmissor, é classificado como uma conexão sináptica excitatória ou inibitória.

As bases para o conexionismo no período de investigação da ciência cognitiva aqui considerado advém da concepção de McCulloch e Pitts (1943), para os quais o sistema nervoso é composto por uma rede de neurônios formados, dentre outras coisas, por um soma e um axônio. O soma do neurônio consiste no seu corpo celular, e o axônio é o cilindro-eixo. As denominadas sinapses são as conexões entre um axônio e a soma de outro neurônio.
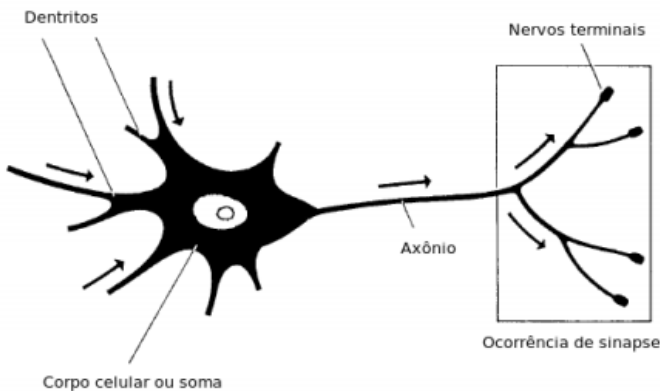
Figura 4. Modelo abstrato do neurônio biológico
Fonte: Adaptado de ARBIB, 2002, p. 4.

Cada neurônio pode receber inúmeros *inputs* de outros neurônios ou do ambiente externo. Entretanto, eles podem produzir apenas uma resposta que é transmitida a outros neurônios ou para o ambiente externo. A resposta do neurônio é enviada pelo axônio através das terminações axonais. As sinapses podem ser divididas entre inibitórias e excitatórias. As sinapses excitatórias auxiliam no disparo, enquanto as inibitórias buscam diminuir a possibilidade do disparo ou mesmo inibi-los totalmente. As conexões entre neurônios no cérebro são feitas a partir da transmissão de substâncias químicas. Os neurônios sempre possuem uma espécie de limiar, cujo estímulo precisa ultrapassar para dar início a um impulso.

A expressão redes neurais artificiais parte da motivação de criar modelos capazes de simular as capacidades do cérebro humano de reconhecer, processar e generalizar dados e padrões. Em geral, esses modelos são utilizados em ambientes nos quais o fluxo de informação muda constantemente. Uma rede neural é um sistema computacional constituído a partir de centenas de unidades básicas que simulam as funções dos neurônios. Esses elementos são interligados, trabalhando em paralelo, para desempenhar uma determinada tarefa. Essas redes de neurônios são responsáveis pelo que chamamos de pensamento, emoção, cognição.

Dentre os primeiros projetos de redes neurais artificiais, McCulloch e Pitts (1943), baseados na ideia do potencial inibitório e excitatório dos neurônios, interpretaram que o seu funcionamento ocorria de maneira semelhante a um circuito binário. Nesse modelo, as conexões entre os neurônios, também chamados de nódulos, são realizadas por meio da transmissão de substâncias eletroquímicas que disparam informações para outros nódulos com base em seu limiar. Estes estímulos, na forma de sinapses, podem ser inibitórios ou excitatórios. Nas RNAs, o limiar e as sinapses são representados a partir de valores numéricos, em vez de substâncias eletroquímicas.

Para McCulloch e Pitts (1943), o disparo dos nódulos pode ocorrer conforme apenas dois estados possíveis, ativado ou desativado. Nas RNAs, o disparo ocorre quando certo valor numérico é atingido pelo nódulo e este valor é determinado por uma fórmula lógica, cujo disparo ocorre apenas se sua fórmula correspondente for verdadeira. O modelo geral dos nódulos parte de algumas pressuposições: $N_i(t)$ significa que um nódulo c_i dispara no tempo *t*. N_i é a *ação* de um neurônio c_i – o tempo é

discreto e pode ser representado por números naturais. Ele é determinado pelas sinapses entre os nódulos: cada sinapse representa um tempo. A *solução* de uma rede *Ñ* é um conjunto de sentenças que regulam o disparo dos nódulos dessa rede. Apresentadas essas pressuposições, podemos ter as seguintes relações:
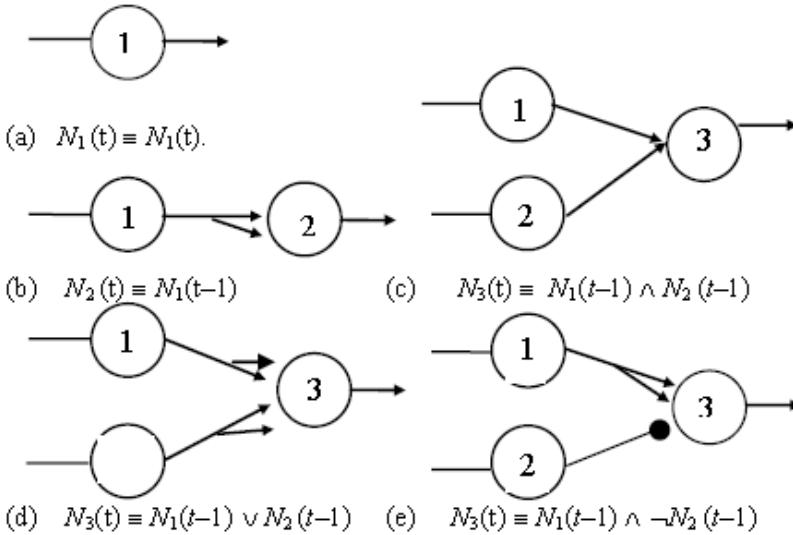
Figura 5. Modelos de redes neurais básicas do sistema apresentado por McCulloch e Pitts.
Fonte: ALVES, 1999, p. 98.

O número no interior de cada nódulo representa a ordem de sua execução, enquanto seu limiar é sempre considerado como sendo 1. Na rede (b) ocorre o que chamamos de demora sináptica. Essas demoras ocorrem, pois a velocidade do impulso produzido pelo disparo dos neurônios pode variar de acordo com o número de sinapses. Assim, expressões com *n* demoras sinápticas, denotadas por "$t - n$", exigem *n* nódulos à esquerda do nódulo calculado para tornar o seu disparo equivalente ao valor de verdade de sua proposição lógica.

A ideia principal defendida pelos pensadores conexionistas, em especial McCulloch (1965), é a de que as RNAs são capazes de simular atividades mentais e nervosas, como, por exemplo, a sensação de frio ou calor. Nos sistemas formais, as RNAs utilizam-se de receptores capazes

de medir a temperatura. Sendo estes receptores adequados e capazes de transmitir a informação para outros nódulos com a capacidade de implicar a sensação de calor ou frio a partir de seu limiar, estipula-se que essas sensações podem ser simuladas por essas redes.

Embora o modelo de McCulloch-Pitts tenha uma construção diferente da máquina de Turing, podemos notar algumas semelhanças entre ambos os modelos. Para computar uma função, a rede deve ser construída de tal maneira que, apresentada alguma entrada, a resposta deve ser correta, caso contrário, deve-se produzir uma rede nova. A máquina de Turing e o modelo McCulloch-Pitts não apresentam regras de aprendizagem e, de certa forma, funcionam de maneira algorítmica.

Na década de 1950 surgiram novas RNAs mais sofisticadas. Uma das grandes vantagens dessas redes é que elas podem reconhecer padrões, mesmo aqueles defeituosos, com algum elemento faltando. Como os nódulos estão interligados com inúmeras camadas de outros nódulos, considera-se que cada um deles também receba parte do padrão total de ativação. Essencialmente, o reconhecimento de padrões para os conexionistas consiste em colocar valor de conexões sinápticas por meio da memorização, pela inserção de informações, por exemplos, por analogia ou por exploração e descoberta. Entretanto, não existe uma definição formal de RNA como existe para a máquina de Turing. Cada rede tem sua particularidade e cada definição geral pode variar de modelo para modelo.

Um dos objetivos fundamentais dos conexionistas posteriores ao modelo de McCulloch-Pitts (1943) era fazer com que as RNAs reconhecessem padrões. Existem dois tipos de padrão: o primeiro é caracterizado, segundo Caudill e Butler (1992), como uma propriedade emergente da dinâmica da interação entre os nódulos de uma rede. Os do segundo tipo são aqueles que as redes são capazes de reconhecer, que são um conjunto de elementos que representam algum objeto. Na simulação de uma rede, um padrão é considerado como um conjunto de valores distribuídos metricamente. Dessa forma, caso desejemos apresentar para uma rede uma letra do alfabeto, a respectiva letra seria representada como um conjunto de valores numéricos e cada nódulo trataria de representar um desses valores do padrão de entrada.

Nem sempre a rede reconhece os padrões na primeira tentativa. A aprendizagem geralmente ocorre a partir da presença ou da ausência

de um elemento capaz de testar o reconhecimento de padrões da rede. Uma realimentação explícita significa que em certos intervalos de tempo um sistema assinala os erros e acertos. No caso em que a realimentação não é explícita, a aprendizagem ocorre sem a presença de um sistema externo. Costuma-se chamar esses dois casos de ensino supervisionado e não supervisionado.

No caso da aprendizagem supervisionada, o sistema externo indica explicitamente o que é considerado um comportamento bom ou um ruim. Por exemplo, imaginemos um caso em que seja desejado o reconhecimento entre os números 1 e 2. Apresentam-se para eles letras sucessivamente aos nódulos de entrada. Observa-se qual dos nódulos de saída parece estar mais excitado. Se for o que convencionou a representar o número que foi apresentado, nada deve ser corrigido, caso contrário deve-se mudar os valores das conexões sinápticas a fim de se chegar ao resultado desejado. No caso da aprendizagem não supervisionada, em vez de informar se a resposta dos nódulos foi correta ou não, usa-se um esquema capaz de induzir a rede a responder de maneira semelhante às regularidades apresentadas. A aprendizagem não supervisionada é marcada pela presença de redes auto-organizadas, ou seja, que possuem a capacidade de criar padrões de comportamento não previsíveis e decentralizados e, em alguns casos, em constante adaptação. Segundo Debrun (2009, p. 54), "[...] uma organização ou 'forma' é auto-organizada quando se produz a si própria".

Aos componentes gerais das RNAs podemos elencar a presença de unidades de processamento (nódulos ou neurônios artificiais), estados de ativação destas unidades (limiar), funções de saída que determinam a resposta de cada nódulo, padrões de conectividade que definem a conexão entre os nódulos, regras de propagação, algum dispositivo capaz de representar o meio e os objetos, e, por fim, de regras que permitam a aprendizagem. Embora apresentemos estas sete características gerais, destacamos que não há uma definição precisa sobre as RNAs, pois cada uma possui as suas próprias particularidades, podendo ou não apresentar grande parte dessas características.

Dentre as redes posteriores à rede construída por McCulloch e Pitts, destacamos o modelo conexionista *perceptron.* Criado na década de 1950 por Rosenblatt (1962), seu criador afirmava que o *perceptron* não devia ser comparado ao modelo de McCulloch-Pitts, pois esse modelo

apresentava um funcionamento quase algorítmico. Além disso, não apresentavam uma regra de treinamento para a aprendizagem da rede.

O *perceptron* foi criado fundamentalmente com o objetivo de modelar a percepção visual. O objetivo para essa rede consiste em classificar padrões em duas classes distintas, A ou B. Os nódulos do *perceptron* funcionam de maneira semelhante ao modelo proposto por McCulloch-Pitts. Se o padrão pertence à classe A, o *perceptron* deve disparar. Se o padrão pertence à classe B, ele não deve disparar A.

Segundo Caudill e Butler (1992), a função transferência do *perceptron* é formada por dois passos: o primeiro consiste no cálculo de entrada total encontrado pela função:

$$I = \sum_{i=1}^{n} w_i x_i$$

Onde são vetores peso e *input*. Ou seja, o *input* total da rede é a somatória da multiplicação de cada valor de entrada pelo seu valor peso.

O segundo passo da função transferência do *perceptron,* segundo Caudill e Butler (1992), é o cálculo da resposta do nódulo, encontrada pela seguinte função:

$$y = \begin{cases} +1, se\ I \geq T \\ -1, se\ I < T \end{cases}$$

Ou seja, o nódulo dispara quando a sua entrada total é maior ou igual ao seu limiar. Caso contrário, não dispara.

O seu treinamento também segue uma respectiva regra:

$$w_{novo} = w_{velho} + \beta_{yx}$$

$$\boldsymbol{\beta} = \begin{cases} +1, \text{se a resposta do } perceptron \text{ está correta} \\ \qquad -1, \text{caso contrário} \end{cases}$$

$$\beta$$

Podemos notar, na função acima, que a mudança de peso necessita da resposta desejada do nódulo. Dizemos, assim, que a sua aprendizagem é supervisionada. O treinamento deve ser feito de maneira organizada e ordenada para que uma rede aprenda a reconhecer um determinado grupo de padrões; para isso, será necessário um algoritmo de treinamento. Segundo Caudill e Butler (1992, p. 29, tradução nossa), no caso do *perceptron*, ele é o seguinte:

> 1. Para cada padrão no conjunto de treinamento
>
> 1.1 aplicar o próximo padrão para o perceptron
>
> 1.2 gravar a resposta do perceptron
>
> 1.3 se a resposta do perceptron está correta,
>
> – e a resposta foi +1, então
>
> o novo vetor peso = velho vetor peso + o vetor do padrão de *input*
>
> – e a resposta foi -1, então
>
> o novo vetor peso = velho vetor peso - o vetor do padrão de *input*
>
> 1.4 se a resposta do perceptron está incorreta,
>
> – e a resposta foi +1, então
>
> o novo vetor peso = velho vetor peso - o vetor do padrão de *input*
>
> – e a resposta foi -1, então
>
> o novo vetor peso = velho vetor peso + o vetor do padrão de *input*
>
> 2. finalize para cada padrão no conjunto de treinamento.

Após efetuar os cálculos para todos os padrões, deve-se analisar se a rede foi capaz de classificar corretamente cada um deles. Se foi capaz de classificar, ela aprende, caso contrário, deve-se recomeçar a fazer os cálculos até que a rede seja capaz de reconhecê-los.

Podemos observar que o conexionismo, semelhante ao cognitivismo, adota como base uma perspectiva funcionalista sobre o estudo da cognição. Dentre seus pressupostos, consta a relevância do uso de modelos, mas, diferentemente do cognitivismo, o conexionismo julga ser importante a construção de modelos análogos à arquitetura dos neurônios biológicos e que a cognição é resultado do processamento em paralelo e distribuído efetuado por vários nódulos. Entre seus quebra-cabeças está a pretensão em criar modelos capazes de simular a cognição, como, por exemplo, as capacidades de reconhecer, processar e generalizar dados e padrões. Nesse sentido, em termos kuhnianos, podemos dizer que a perspectiva conexionista, embora possua as suas diferenças em relação ao cognitivismo, parece não possuir uma diferença propriamente paradigmática.

O comportamento inteligente para os conexionistas está associado fundamentalmente à capacidade de aprender a reconhecer padrões, espécie de representação de algum objeto. Por exemplo, um limão pode ser representado por diversas características, como de ser fruta, ser verde, ser azedo etc. Tais características formam a ideia de limão. Reconhecemos este padrão quando somos capazes de classificar o objeto representado pelo padrão de acordo com sua classe.

Cada elemento do padrão é transformado em um determinado número. Na grande maioria das vezes, no processo de simulação das redes, os padrões transformam-se em conjunto de números. A rede reconhecerá o padrão quando seus nódulos de saída dispararem de tal modo que o classifique corretamente.

Para os conexionistas, a aprendizagem nas redes neurais artificiais acontece fundamentalmente através do ajuste dos pesos da rede, ou seja, por meio de um processo de treinamento. Uma vez ajustados os pesos, como afirmam Caudill e Butler (1992, p. 9, tradução nossa), "[...] pode-se dizer que a rede aprendeu." Desse modo, a aprendizagem não tem como princípio fundamental o seguimento de regras e manipulação de símbolos. Diferentemente, ela acontece através do fortalecimento das conexões entre nódulos.

A ideia de analisar a aprendizagem através da alteração das conexões entre os neurônios teve como um de seus primeiros adeptos o pesquisador Hebb (1949), que estabeleceu uma lei, conhecida como lei

de Hebb. Ela basicamente diz o seguinte: quando um neurônio estimula outro de tal modo que o faça disparar, a conexão a partir da primeira célula com a segunda é fortalecida (Hebb, 1949).

Por ser capaz de realizar comportamentos inteligentes, via reconhecimento de padrões, o projeto conexionista começou com muita euforia na década de 1950. Tal euforia, porém, durou pouco tempo. Minsky e Papert (1969) mostraram que o *perceptron* não podia resolver problemas não linearmente separáveis. Um dos exemplos desses problemas é o do ou-exclusivo da lógica proposicional. Esses dois pesquisadores previram ainda que nem *perceptrons* mais sofisticados com camadas intermediárias seriam capazes de resolvê-los. Isso porque não haveria um bom modo de saber quais são as respostas desejadas para os nódulos das camadas intermediárias da rede. Além de não poder resolver problemas linearmente inseparáveis, Minsky e Papert (1969) mostraram que o *perceptron* não é capaz de separar mais do que duas classes distintas.

Outra forte crítica direcionada à hipótese de que os modelos conexionistas podem ser considerados inteligentes é feita por Dreyfus (1993, p. xxxviii, tradução nossa), para o qual "[...] a rede exibirá a inteligência nela embutida pelo projetista para aquele contexto, mas não terá o senso comum que lhe permitiria adaptar-se a outros contextos". O autor chega a essa conclusão por meio da análise de experiências feitas com redes para reconhecimento de padrões. Um desses trabalhos consistia em fazer com que uma rede neural artificial reconhecesse a presença de tanques de guerra numa floresta. Segundo Dreyfus (1993, p. xxxvi, tradução nossa), o exército tirou um

> [...] certo número de fotografias de uma floresta sem os tanques e, em seguida, alguns dias mais tarde, com os tanques aparecendo claramente por trás das árvores, e treinaram uma rede para que ela distinguisse os dois tipos de fotos. Os resultados foram impressionantes, e o exército ficou ainda mais impressionado quando se soube que a rede podia generalizar seu conhecimento para fotos que não haviam feito parte do conjunto de treinamento. Só para se ter certeza de que a rede estava de fato reconhecendo tanques parcialmente ocultos, no entanto, os pesquisadores tiraram mais fotos na mesma floresta e as mostraram à rede treinada.

Porém, neste novo lote de fotos, a rede não obteve sucesso. Não conseguiu distinguir fotos com tanques parcialmente escondidos atrás de árvores e fotos sem nada atrás delas. O que se descobriu depois é que o primeiro lote de fotos havia sido tirado em dias distintos, um ensolarado e outro não. Desse modo, a rede não aprendeu a reconhecer a existência de tanques na floresta, mas, sim, aprendera a reconhecer florestas com e sem sombras. Dreyfus (1993, p. xxxviii, tradução nossa) afirma que

> [...] os projetistas de redes não mais podiam permitir que seus sistemas fossem "treinados" sem pré-especificar e, portanto, restringir, a classe de generalizações apropriadas permitida para o problema (ou "espaço de hipótese"). A arquitetura das redes é assim planejada para transformar *inputs* em *outputs* "somente das maneiras que estão no espaço de hipótese".

Entretanto, é claro que um ser humano, por exemplo, seria capaz de reconhecer a existência de tanques em muitos contextos distintos. Não há necessidade de restringir o cenário a um específico, ou sempre determiná-lo.

As dificuldades no projeto conexionista fizeram com que ele tivesse seu progresso abalado por um longo tempo. Abalado, mas não eliminado. Alguns pesquisadores, como Hopfield (1982), Kohonen (1987), Rumelhart e McClelland (1986), continuaram trabalhando firmes em seus propósitos conexionistas. Na década de 1980, criaram redes capazes de resolver problemas não linearmente separáveis e reconhecer objetos, dividindo-os em mais do que apenas duas classes distintas. Eles concordavam com o argumento de Minsky e Papert (1969) de que as redes de multicamadas não eram capazes de resolver esses problemas. Isso se a regra de treinamento fosse a mesma do *perceptron* simples. Porém, se essa regra fosse modificada ou aperfeiçoada, tal argumento seria refutado. Foi o que fizeram os conexionistas: criaram redes cuja regra de treinamento é uma extensão da regra do *perceptron* simples, denominada *backpropagation*. Dentre as redes com tal regra, encontramos o *perceptron* multicamadas e as redes de Kohonen. Os sucessos, alcances e limites dessas novas redes e seu poder explicativo, no entanto, fogem ao escopo deste trabalho, embora sejam ilustrativos da tentativa dos adeptos de uma matriz curricular em

tentar firme e insistentemente resolver seus quebra-cabeças e em não abandonar seu paradigma.

Como procuramos mostrar, o cognitivismo e o conexionismo eram as duas principais vertentes da jovem ciência cognitiva. Seriam essas posições duas candidatas a paradigmas, seriam teorias distintas de um paradigma apenas ou nenhuma dessas alternativas? Sobre isso tratamos a seguir, a fim de encerrar nossa análise neste trabalho.

Considerações Finais
O estatuto científico da ciência cognitiva em sua fase inicial

Na tentativa de explicar o funcionamento da ciência, Thomas Kuhn (2011a) desenvolve a sua estrutura das revoluções científicas. Levando em consideração a história da própria ciência e adotando uma postura falibilista das teorias científicas, empiricamente testáveis e envoltas em uma matriz disciplinar, este pensador contemporâneo propõe uma abordagem estruturalista da atividade e constituição dessa atividade humana. Com isso, se propõe a explicitar o processo que permite o seu progresso, ou seja, o aprimoramento de nosso conhecimento do mundo.

Nessa abordagem, o progresso pode se dar em dois sentidos: um no sentido geral, outro em um sentido restrito. O primeiro deles, o que se refere ao progresso em um sentido amplo, é caracterizado pela mudança paradigmática radical. É amplo tanto pela mudança em uma área de pesquisa tanto quanto na ciência como um todo. Uma vez que as áreas de pesquisa estão conectadas, a mudança em uma área de pesquisa também acarreta mudanças nas outras áreas. Por isso, nesse sentido, uma revolução científica é uma mudança geral, que acarreta, dentre outras coisas, uma visão de mundo distinta, que exige novas metodologias, métodos, pressupostos metafísicos, dentre muitos outros elementos. A mudança paradigmática de geocentrismo para o heliocentrismo, por exemplo, não provocou uma alteração apenas na astronomia, mas também na física, na medicina, favoreceu o surgimento de outras áreas como a química. Tal

revolução foi provocada por modificações em outras áreas de pesquisa e também implicou alterações nelas, em uma causalidade circular. Um dos elementos metodológicos, ontológicos e epistemológicos alterados no espírito científico da época com tal ruptura foi o fato de que a atenção não era mais às qualidades dos objetos ou fenômenos, como na ciência de origem aristotélico-tomista, mas sim à quantidade. Assim, não importa mais o lugar, o tamanho ou cor do objeto, mas sim quanto ele mede, pesa, a distância em que se encontra de um dado ponto e assim por diante. Nesse sentido, a matemática passou a ser uma ferramenta metodológica fundamental na ciência moderna.

Considerando que a ciência é uma prática humana, a própria humanidade sofre grandes mudanças estruturais nas revoluções científicas. No caso da revolução copernicana ilustrado acima, o ser humano foi retirado do centro do universo, seja ele próprio, seja o lugar que habita. Já no século XVIII, a abordagem darwiniana, a partir das ciências biológicas, buscou mostrar, dentre outras coisas, que somos membros do reino animal, sem qualquer superioridade ou privilégios com respeito aos demais seres. Um dos elementos metafísicos fundamentais nessa abordagem foi o pressuposto de que o mundo não está organizado e ordenado de tal forma devido a uma inteligência suprema. Ao contrário, a inteligência ou a consciência em certos sistemas é oriunda de partes não inteligentes e não conscientes que se unem, muitas vezes ao acaso, produzindo a inteligência e a consciência. Assim, o plano de pesquisa científico passou a ser buscar uma explicação de como ocorre tal transformação, o que, de alguma forma é o procedimento na ciência cognitiva: partes integradas, sejam instruções de um programa, sejam neurônios, em si, não inteligentes ou conscientes, se relacionam e produzem ações inteligentes ou processos cognitivos conscientes em geral.

Pouco depois de Darwin, a psicanálise freudiana sugeriu que nossas mentes estão repletas de mecanismos inconscientes. Apesar das indagações a respeito da cientificidade dessa área de pesquisa, tais propostas resultaram em mudanças paradigmáticas, mostrando que, além de não sermos o centro do universo, tampouco somos preponderantes às outras espécies ou possuímos completo controle sobre nossa mente. Elas provocaram uma espécie de processo de deslocamento e reavaliação da natureza fundamental da humanidade e o seu papel no meio ambiente. Isso exigiu novas configurações em matrizes curriculares de diversas áreas

de pesquisa e na própria ciência, investigando, inclusive, se elas poderiam ser consideradas científicas.

No século XX, outra grande mudança de perspectiva, ainda em curso, é a virada informacional. Ela substitui a visão de mundo centrada na mente humana para a explicação da realidade, da cognição e da própria natureza humana, atribuindo papel de destaque à informação nesses estudos. A virada informacional teve seu surgimento fortalecido por dois aspectos de alto impacto, na ciência contemporânea: o teste de Turing e a segunda lei da termodinâmica. Como exposto anteriormente, Turing (1950) propõe uma definição de pensamento enquanto processamento de informação, construindo uma abordagem, de alguma forma, empírica, em detrimento das concepções metafísicas a respeito da cognição. Tal definição, inclusive, propicia ou permite considerar as teorias ou hipóteses da ciência cognitiva empiricamente testáveis, pelo menos no cognitivismo. No caso do conexionismo, o mesmo poderia ser dito a partir dos pressupostos estabelecidos por McCulloch (1965). Esses dois elementos são de suma importância, necessários, mas não suficientes, para poder considerar as abordagens da ciência cognitiva como paradigmas. A termodinâmica, por sua vez, além de inserir a informação no contexto teórico da física, conforme mostra Mitchel (2009), também contribuiu, a partir do uso de conceitos, como os de ordem e entropia, como fonte de inspiração para a concepção quantitativa da informação.

Como procuramos mostrar neste livro, é nesse contexto que nasce a ciência cognitiva. Por um lado, seu surgimento foi favorecido e implicado por tais mudanças de visão de mundo. Por outro lado, estas mudanças foram resultado de estudos em áreas como a cibernética que, de uma forma ou de outra, resultaram na emergência da ciência cognitiva. Assim, em um sentido geral, essa nova área foi influenciada e influenciou um novo ciclo de atividade científica, um novo modelo de ciência. Em tese, tais reconfigurações permitem o aprimoramento do conhecimento científico sem, entretanto, adotar uma perspectiva linear de progresso, no sentido de aproximação da verdade, como propunham os adeptos do Círculo de Viena, ou empiristas lógicos, por exemplo.

Kuhn (2011a) visa a tratar, em sua abordagem, desse modo de revolução, explicativo do progresso científico, que não ocorre de

um instante para outro, mas pode ser um processo secular. Como procuramos mostrar, é em um desses processos revolucionários que surge a ciência cognitiva.

Mas há também outro sentido importante de ampliação de conhecimento em uma área de pesquisa, que consiste no aprimoramento do seu paradigma dominante. Tal progresso interno, possível enquanto ciência normal, consiste, dentre outras coisas, na descoberta de novas técnicas de pesquisa e de experimentos, de resoluções de quebra-cabeças ou de anomalias, estabelecimento ou fortalecimento de conceitos, pressupostos metafísicos e metodológicos em uma área de pesquisa. Esse desenvolvimento da matriz disciplinar é possível graças à convergência da comunidade científica em torno de um único paradigma. Trata-se de um momento de calmaria, no qual, em geral, não há espaço para críticas ou dúvidas sobre o poder explicativo e preditivo de teorias, da segurança oferecida pelo paradigma como um norte para a prática científica. Não se trata de uma dúvida dogmática no sentido de crença ou fé inabalável, mas pela confiança advinda de resultados já oferecidos pelo paradigma na resolução de quebra-cabeças.

É nesse segundo sentido de progresso que este livro busca analisar o estatuto da ciência cognitiva. Para tanto, necessitamos indagar se ela já começa no estágio de ciência normal ou se surge como a maioria das áreas de pesquisa, em um período de pré-ciência, de luta paradigmática.

Com base no que apresentamos anteriormente, podemos obter alguns elementos indicativos da situação dessa área de pesquisa em sua fase inicial. De algum modo, a questão, em termos kuhnianos, seria responder se havia um paradigma dominante, se é que havia, ou percebemos a existência de paradigmas rivais disputando a hegemonia. Podemos buscar um indício de resposta a essa questão considerando o contexto histórico e o contexto epistemológico, tal como desenvolvidos nos dois últimos capítulos deste livro.

Conforme observado por Kuhn, são poucas as disciplinas que começam como ciência normal. Em geral, esses casos se configuram como dissidências ou ramificações de outras áreas já existentes, em geral com especificações no objeto de pesquisa ou fenômenos do mundo analisados.

Considerando este rápido histórico do princípio da ciência cognitiva exposto anteriormente, a ciência cognitiva, de uma forma ou de outra, teve o surgimento influenciado pela cibernética. A ciência do controle possuía princípios, métodos, regras metodológicas, uma base metafísica e outros elementos paradigmáticos adotados para o estudo de seu objeto de estudos, em particular, dos processos cognitivos. Parece ter havido um esforço de consolidação de um paradigma nessa área. Nas conferências descritas no segundo capítulo podemos perceber o esforço em fortalecer e delimitar conceitos, buscando oferecer uma série de princípios metodológicos e conceituais coesos. Técnicas e procedimentos para o estudo da percepção, da memória, analisando como ocorre o armazenamento e processamento de informação pelo cérebro foram apresentados e discutidos nesses encontros.

As Conferências Macy buscaram, ainda, constituir uma comunidade científica coesa em torno de uma matriz disciplinar. Foram estabelecidos pressupostos metodológicos tais como o uso de modelos no estudo da cognição e a pesquisa interdisciplinar, uma certa preferência pelo fisicalismo e pela teoria do controle. O conceito de teleologia surge como um princípio metodológico capaz de oferecer subsídios à manipulação e à compreensão da natureza, a partir de um escopo teórico específico.

Em seus encontros, os pesquisadores propunham conceitos e buscavam consolidá-los, na tentativa de criar uma unidade conceitual. Ofereciam hipóteses explicativas testáveis empiricamente para certos fenômenos cognitivos envolvendo ou sobre a memória, a percepção, a linguagem, as emoções, a ação. Havia o estabelecimento de uma agenda científica, que assegurava quais fenômenos deveriam ser considerados e como eles deveriam ser abordados. A linguagem e a criação de conceitos aparecem como elementos fundamentais na abordagem de Kuhn para a consolidação de uma comunidade científica. Tais elementos implicam uma descrição do mundo e uma série de pressupostos teóricos que auxiliam na atividade científica dos membros da comunidade. Faz parte do conceito de paradigma toda uma constelação de técnicas e visões de mundo.

Ademais, surgiram revistas de divulgação científica e manuais explicativos em que neófitos poderiam se iniciar na atividade científica na área. Assim, o conjunto de adeptos ao novo modelo de pesquisa não

contaria mais apenas com um conjunto restrito de pesquisadores, tampouco seria apenas uma área com um conjunto de intenções ou promessas de pesquisa e de descobertas. Ao contrário, o objetivo de seus adeptos iniciais era expandir a comunidade, vislumbrando constituir uma coletividade inclinada, direcionada e motivada em montar os quebra-cabeças da área.

O contexto histórico exposto aqui, portanto, parece indicar, em um primeiro olhar, um avanço nos estudos sobre a cognição, a consolidação de um arcabouço conceitual e uma comunidade científica sólida em busca da explicação de certos fenômenos. Desse modo, a conclusão seria, em princípio, de que a ciência cognitiva é uma das poucas áreas a se constituir, desde o seu início, como uma ciência normal, experimentando, de imediato, um período de progresso interno.

Entretanto, uma análise mais cuidadosa, a partir do breve relato histórico por nós exposto, mostra que, antes desta calmaria, a nova área passava por fortes tempestades em um navio frágil e sem muita segurança. Graças a desavenças internas do grupo de pesquisadores reunidos em torno do estudo dos processos cognitivos e a um momento histórico de grande efervescência, observamos que o sonhado paradigma cibernético não chegou a se concretizar tal como esperado ou desejado por seus defensores, tais como Wiener.

As Conferências Macy mostram, de alguma maneira, a existência de inúmeras dificuldades e empecilhos para a possibilidade de acordo entre seus participantes. Em primeiro lugar, os dois textos assumidos como basilares para o surgimento da ciência cognitiva já ilustram uma diferença essencial entre seus integrantes. O próprio nome de ambos os artigos já revela tais diferenças: "Behavior, Purpose and Teleology" e "A logical calculus of the ideas immanent in nervous activity". Por um lado, o grupo de Rosenblueth, Wiener e Bigelow estava tão inclinado aos anseios da cibernética que deu origem à versão cognitivista, exposta no terceiro capítulo. Por outro lado, a turma de McCulloch e Walter Pitts defendia uma versão não sistêmica, originária do conexionismo, igualmente exposto no terceiro capítulo.

Embora já tenhamos discutido algo sobre o assunto, poderíamos dizer, especificamente, que, em termos epistemológicos, tal como exposto no terceiro capítulo, a ciência cognitiva tampouco iniciou como

ciência normal. Tal como exposto acima sobre o aspecto histórico, no contexto epistemológico também, em princípio, parecia haver uma base constituinte de um paradigma. Havia, por exemplo, uma agenda científica, com fenômenos aparentemente bem delimitados a serem investigados: certo tipo de processos cognitivos, levando em consideração alguns aspectos destes processos e ignorando outros. Havia uma definição de algum modo consensual do que se entendia por cognição. Processos cognitivos envolvem raciocínio, percepção, linguagem, conhecimento, aprendizagem, inteligência. Havia um conjunto de conceitos estabelecidos. Dentre eles, alguns fundamentais, como o de que a mente é um sistema de processamento de informação, não uma substância imaterial, não física. A inteligência era entendida como processamento eficiente de informação, ao estilo do teste de Turing.

Como mostramos na segunda seção do terceiro capítulo, havia uma série de pressupostos metodológicos, a inclinação por estudos interdisciplinares e, principalmente, a defesa de uma perspectiva representacionista e do uso de modelos nos estudos da cognição. No entanto, como também já exposto acima, as divergências pareciam ser muito maiores ou pelo menos tão relevantes quanto as convergências entre os grupos de pesquisadores.

Apesar de concordar que a representação mental apresenta um papel fundamental nos estudos cognitivos e na própria cognição, as duas correntes expostas no terceiro capítulo as entendem de modo radicalmente distinto. Cognitivistas defendiam e se baseavam na concepção de representações simbólicas, de processamento algorítmico e sequencial, de funcionamento mental ou cognitivo ao estilo máquinas de Turing. Já os conexionistas entendiam as representações como relações subsimbólicas entre elementos físicos de uma rede, auto-organizada, cujo processamento ocorre de modo distribuído e paralelo, passíveis de modelagem por sistemas do tipo redes neurais artificiais.

Seria possível dizer que estas duas abordagens poderiam ser consideradas duas teorias explicativas de fenômenos envoltas em um mesmo paradigma, como ilustramos na terceira seção do primeiro capítulo sobre a matéria escura na física? Acreditamos que este não é o caso. Em primeiro lugar, a resposta negativa se deve aos princípios metafísicos

envolvidos em ambas as vertentes. Embora concordem em alguns pontos, havia duas visões de mundo distintas nelas. A primeira turma parecia mais atraída pelo funcionalismo, enquanto a segunda estava mais inclinada pela aceitação de um fisicalismo, ou de uma teoria da identidade mente-cérebro. Embora o funcionalismo não seja contraditório ao fisicalismo, essa diferença ideológica já representaria visões de mundo incompatíveis para constituir um único paradigma.

As diferenças metafísicas produzem resultados na própria pratica científica. No caso de Rosenblueth e sua turma, os modelos considerados adequados para o estudo da cognição seriam as máquinas do tipo Turing, que manipulam símbolos seguindo regras lógicas, de modo que pensar significa calcular. Já os adeptos da proposta de McCulloch adotariam como modelo para o estudo da cognição sistemas semelhantes ao cérebro, que processam informação de modo não linear, formal, simbólico, algorítmico. Desse modo, de alguma forma, ambos os grupos, apesar de estudarem os mesmos fenômenos, os entendem e interpretam de modo radicalmente distinto.

As próprias Conferências Macy ilustram embates radicais entre seus participantes envolvendo divergências conceituais, metodológicas, teóricas. Nesse sentido, do ponto de vista histórico, a ciência cognitiva parece não ter começado como ciência normal. Na melhor das hipóteses, podemos dizer que ela começa como a grande maioria das áreas de pesquisa, ou seja, por um período de disputa paradigmática.

Em um sentido mais radical, poderíamos dizer que não havia disputa entre paradigmas, uma vez que nem sequer havia paradigmas disputando naquele momento. De alguma maneira, os elementos constituintes de uma matriz disciplinar ainda estavam sendo formados. Esse era o objetivo das Conferências Macy. No entanto, como tentamos mostrar, havia elementos que pelo menos indicavam matrizes disciplinares em formação. Esses elementos nos indicam que, se considerarmos esse período como base para a ciência cognitiva, este seria já o seu período de pré-ciência, pelo menos em um sentido ampliado do termo, uma vez que, formalmente, a ciência cognitiva foi estabelecida alguns anos mais tarde, como mostramos no terceiro capítulo.

O fato de não ter começado como ciência normal retardou o seu desenvolvimento, pelo menos em termos epistemológicos. As diferenças entre seus integrantes exigiram a construção de um corpo de conceitos, princípios e outros fatores para somente depois começar a criar teorias explicativas dos fenômenos investigados.

Ademais, diversas anomalias foram expostas às duas principais correntes naquele momento. Ao cognitivismo, foram aplicadas, por exemplo, os resultados dos teoremas da incompletude de Gödel e críticas semelhantes, como mostrado no terceiro capítulo, especialmente o argumento do quarto chinês de Searle ou as críticas de Penrose. Elas buscavam mostrar que sistemas formais não eram capazes de oferecer modelos explicativos da cognição e sequer simular aspectos cognitivos. Nesse bojo estavam também as críticas de que máquinas do tipo Turing não aprendem, uma vez que funcionam seguindo regras lógicas, manipulando símbolos, sem a capacidade de adaptação ou correção de erros. Os cognitivistas se defendem dizendo que suas máquinas, principalmente as probabilísticas, como os programas de jogos de xadrez, por exemplo, são capazes de corrigir erros. O próprio Turing, em seu artigo de 1950, expressa como a máquina é capaz de surpreendê-lo tomando cursos de ação impensados ou imprevistos por ele.

Outra crítica dirigida aos modelos cognitivistas era o fato de não serem bons no reconhecimento de padrão. Ou seja, tais sistemas não eram bons modelos para a explicação de processos cognitivos que envolvem a cognição. Isso não apenas pela limitação de tais modelos, mas porque a percepção não funciona como tais modelos. Não se trata apenas de uma questão heurística ou metodológica. Trata-se de uma questão ontológica.

Outro ponto crítico do cognitivismo, relacionado ao da percepção, é o referente ao conhecimento não proposicional. Máquinas do tipo Turing são boas para calcular, para jogar xadrez, para fazer operações matemáticas. No entanto, não explicam habilidades que envolvem a intuição, adaptação, percepção direta, em atividades como andar de bicicleta, por exemplo.

Diante dessas dificuldades do cognitivismo, poderíamos imaginar que o conexionismo estaria em vantagem epistemológica em relação aos seus opositores. Entretanto, diversas anomalias também eram apontadas a essa versão da ciência cognitiva. Dentre elas estão as já explicitadas no

final do terceiro capítulo, especialmente por Minsky e Papert em relação às redes neurais artificiais da época. Ironicamente, os modelos alternativos às máquinas do tipo Turing para simular aspectos cognitivos referentes à percepção, por exemplo, eram incapazes de apresentar bons resultados justamente no reconhecimento de padrões.

Esse momentâneo fracasso das redes neurais foi um dos motivos pelos quais o conexionismo ficou enfraquecido em comparação ao cognitivismo. A criação de máquinas que jogam xadrez com eficiência, fazem cálculos, resolvem teoremas, processam informações de modo muito mais eficiente e rápido que o ser humano favoreceram a prevalência do projeto cognitivista em detrimento do conexionista. Este consegue se reerguer apenas na década de 1980. Tais motivos influenciaram pesquisadores a aderir ao projeto cognitivista, fazendo jus ao que diz Kuhn sobre a escolha de paradigmas, conforme exposto no primeiro capítulo.

À parte as críticas expostas a cada uma das duas principais vertentes da ciência cognitiva em sua fase inicial, foram despontando, aos poucos, outras correntes de pensamento investigativo nessa área. Uma delas é a Cognição Situada e Incorporada. Dentre seus principais defensores, encontramos Varela, Maturana e Flores (1980), para os quais a cognição, em suas estruturas, emerge de esquemas sensório-motores vivenciados que permitem que a ação seja construída e orientada pela percepção. É a estrutura experiencial sensório-motora contextualizada, a maneira pela qual o sujeito que percebe estar inscrito em um corpo, que determina como o sujeito pode agir e ser modulado pelos eventos do ambiente.

Enquanto as duas correntes expostas aqui defendem, cada uma a seu modo, que a inteligência consiste na capacidade de resolução de problemas previamente especificados, a cognição situada e incorporada entende que os problemas são construídos no próprio processo cognitivo. Ademais, essa perspectiva mostra muitas situações em que a cognição não é mediada por representação, mas ocorre por meios não representacionais, pelo processo de enação. Ao contrário das outras duas correntes, que entendem a cognição como um processo mental, de manipulação de informações, a cognição situada e incorporada, como o próprio nome indica, entende que a cognição envolve e não pode ser entendida sem uma perspectiva sistêmica, na qual são fatores essenciais o meio ambiente, o

corpo e elementos históricos, intimamente relacionados entre si e que não podem ser estudados separadamente.

Em suma, além da ciência cognitiva não ter conseguido se estabelecer como uma ciência normal em seu surgimento, parece ter encontrado ao longo de seu percurso novos projetos de matrizes disciplinares buscando um lugar ao sol. Por um lado, isso mostra que essa área de pesquisa ainda não estava pronta para se estabelecer como uma ciência madura. Por outro lado, muitos resultados foram oriundos dessa atividade cientifica. Podemos dizer que, de modo bastante direto, os atuais robôs ou demais sistemas artificiais, sejam físicos, sejam virtuais, são oriundos desses projetos.

O atrativo tecnológico, seja por influência econômico-financeira, por questões sociais ou psicológicas, favoreceu pesquisas referentes à produção de máquinas de variados tipos, em particular os humanoides. Embora possua elementos epistemológicos, de entender em que consiste a cognição, relacionada à ação, e emoções, muitos projetos de construção de sistemas processadores de informação enfocam a simulação de aspectos cognitivos em sistemas de inteligência artificial, em detrimento do conhecimento e propostas teóricas desses fenômenos. Por um lado, essa escolha se deve aos anseios dos financiadores das pesquisas nessa área. Por outro lado, é consequência do enorme desafio da montagem do quebra-cabeças da cognição.

É necessário, por exemplo, um conhecimento mais aprofundado do cérebro humano, das conexões entre seus elementos, de suas relações com o corpo e com o ambiente. É preciso um entendimento mais aprofundado de como emerge a consciência, de como tratar elementos subjetivos de uma forma objetiva, ou seja, de elementos de primeira pessoa sendo tratados em uma perspectiva de terceira pessoa, ou pelo menos de modo intersubjetivo; descobrir quais elementos químicos, físicos, elétricos, informacionais seriam responsáveis pelo surgimento ou realização de processos cognitivos em um sistema e se eles seriam necessários para a sua existência; avaliar em que medida elementos culturais possuem conexão e relações com o surgimento e existência de elementos cognitivos, bem como sua relação com emoções, sentimentos e ação.

A lista de peças a serem juntadas poderia ser aumentada sobremaneira aqui. Provavelmente a montagem do jogo ainda se delongará por muito tempo. Além da quantidade de unidades a serem conectadas, há muitas faltando, necessitando ser juntadas ao tabuleiro. Talvez estejamos, inclusive, jogando com as peças erradas. Seguindo os conselhos de Kuhn, o que podemos fazer é nos manter motivados em busca da solução do problema, confiantes de que algum dia conseguiremos montar pelo menos algumas partes do desenho cognitivo.

Referências

ALVES, M. A. **Mecanismo e Inteligência**: Um estudo sobre o conceito de Inteligência na ciência cognitiva. 1999. 201 f. Dissertação (Mestrado em Filosofia) – Universidade Estadual Paulista, Faculdade de Filosofia e ciências; Marília 1999.

ALVES, M. A. Reflexões acerca da natureza da ciência: comparações entre Kuhn, Popper e Empirismo Lógico. **Kínesis**, Marília, v. 5, p. 193-211, 2013.

ALVES, M.A; VALENTE, A.R. A estrutura das revoluções científicas de Kuhn: uma breve exposição. **Griot**: revista de filosofia. Vol. 20, n. 1. p. 173-192, 2020.

ARBIB, M. A. **The Handbook Of Brain Theory and Neural Networks**. Londres: The MIT Press, 2002.

BODEN, M. **The Philosophy of Artificial Intelligence**. New York: Oxford University Press, 1990.

BORRADORI, G. **A filosofia americana**: Conversações com Quine, Davidson, Putnam, Nozick, Danto, Rorty, Cavell, MacIntyre e Kuhn. Tradução de Álvaro Lorencini. São Paulo: Unesp, 2003.

CAUDIL, M; BUTLER, C. **Understanding neural networks**: computer explorations. Massachusetts: MIT Press, 1992.

CAUDILL, M; BUTLER, C. **Understanding neural networks**: computer explorations. Massachusetts: MIT Press, 1992. 2 v.

CHARNIAK E. **Naturally intelligent systems**. Massachusetts: MIT Press, 1993.

CHARNIAK, E; McDERMOTT, D. **Introduction to artificial intelligence**. Addisson–Wesley, 1985.

CHOMSKY, N. **The Logical Structure of Linguistic Theory**. 1955. Tese (Doutorado) – University of Pennsylvania, Philadelphia-PA, 1955. Publicado como monografia: Nova York: Plenum Press, 1975.

CHURCHLAND, P. **Matter and consciousness**. Cambridge: MIT Press, 1984.

CORBETT, T. *et al.* Constraining anomalous Higgs boson interactions. **Phys. Rev**. 86, 2012

DAWSON, M. **From Embodied Cognitive Science to Synthetic Psychology**. Proceedings of the first IEEE International Conference on Cognitive Informatics, 2002.

DEBRUN, M. Identidade Nacional Brasileira e Auto-organização. In: D'OTTAVIANO, I. M. L.; GONZALEZ, M. E. Q. (Org.). **Auto-organização nas ciências Exatas e Naturais**. Campinas: Unicamp, 2009.

DREYFUS, H. **What computers still can't to do**. New York: Harper & Row, 1993.

DUPUY, J. P. **Nas origens das ciências cognitivas**. Tradução de Roberto Leal Ferreira. São Paulo: Unesp, 1996.

FEIGENBAUM, E; FELDMAN, J. **Computers and thought**. New York: McGraw–Hill Book Company, 1968.

FODOR, J.A. What psychological states are not. In: BLOCK, N. (Org.). **Readings in philosophy of psychology**. v. 1. Cambridge: Harvard University Press, 1980. p. 237-250.

FRAWLEY, W. **Vygotsky e a ciência cognitiva**: linguagem e interação das mentes social e computacional. Tradução de Marcos A. G. Domingues. Porto Alegre: Artes Médicas Sul, 2000.

GARDNER, H. **A nova ciência da mente**: uma história da revolução cognitiva. Tradução de Claúdia Malbergier Caon; prefácio de Marcos Barbosa de Oliveira. São Paulo: Edusp, 1996.

HANSON, N. Observação e Interpretação. In: **Filosofia da ciência**. Tradução de Leonidas Hegenberg, Octany Silveira da Mota. São Paulo: Cultrix, 1975.

HASELAGER, W. F. G. O mal estar do representacionismo: sete dores de cabeça da ciência cognitiva. In: FERREIRA, A.; Gonzalez, E. Q.; COELHO, J. G. (Ed.). **Encontros com as ciências cognitivas**, v. 4. São Paulo: Coleção Estudos Cognitivos, 2004. p. 105-120.

HEBB, D. **Organization of behavior**. New York: John Wiley, 1949.

HOPFIELD, J. Neural networks and physical systems with emergent collective computational abilities. **Proc. Natl. Acad. Sci**. v. 79, p. 2554-2558, 1982.

HOYNINGEN-HUENE, P. **Reconstructing scientific revolutions**. Chicago: The University of Chicago Press, 1993.

HUME, D. **Investigações sobre o entendimento humano e sobre os princípios da moral**. Tradução de José Oscar de Almeida Marques. São Paulo: Unesp, 2004.

KOHONEN, T. **Self organization and associative memory**. Berlin: Springer-Verlag, 1987.

KOVÁCS, Z. L. **Redes neurais artificiais***:* fundamentos e aplicações. São Paulo: Livraria da Física, 2006.

KUHN, T. **A estrutura das revoluções científicas**. Tradução de Beatriz Vianna Boeira e Nelson Boeira. São Paulo: Perspectiva, 2011a.

KUHN, T. **A tensão essencial***:* estudos selecionados sobre tradição e mudanças científicas. Tradução de Marcelo Amaral Penna-Forte. São Paulo: Unesp, 2011b.

LACEY, H. **Valores e atividade científica**. São Paulo: Discurso Editorial, 1998.

LAKATOS, I. **La metodología de los programas de investigación científica**. Madrid: Alianza, 1989.

LAKATOS, I. O falseamento e a metodologia dos programas de pesquisa científica. In: LAKATOS, I; MUSGRAVE, A. (Org.). **A crítica e o desenvolvimento do conhecimento**. São Paulo: Cultrix, 1979.

LAUDAN, L. **Science and values**: the aims of science and their role in scientific debate. Berkeley: University of California Press, 1984.

LURIA, A.R. **Desenvolvimento cognitivo**: seus fundamentos culturais e sociais. Tradução de L. M. Barreto, M. K. Oliveira, M. M. M. de Andrade, e R. H. Maciel. São Paulo: Ícone, 1990.

MACY 6. **Cybernetics***:* circular causal and feedback mechanisms in biological and social systems. Transactions of the sixth conference, March 24-25, 1949, New York, N. Y. Editado por Heinz von Förster. Josiah Macy Jr. Foundation, New York, 1950, 209p.

MACY 7. **Cybernetics***:* circular causal and feedback mechanisms in biological and social systems. Transactions of the seventh conference, March 23-24, 1950, New York, N. Y. Editado por Heinz von Förster. Editores Assistentes: Margareth Mead e Hans Lukas Teuber. Josiah Macy Jr. Foundation, New York, 1951, 251p.

MACY 8. **Cybernetics**: circular causal and feedback mechanisms in biological and social systems. Transactions of the eighth conference, March 15-16, 1951, New York, N. Y. Editado por Heinz von Förster. Editores Assistentes: Margareth Mead e Hans Lukas Teuber. Josiah Macy Jr. Foundation, New York, 1952, 240p.

MACY 9. **Cybernetics**: circular causal and feedback mechanisms in biological and social systems. Transactions of the ninth conference, March 20-21, 1952, New York, N. Y. Editado por Heinz von Förster. Editores Assistentes: Margareth Mead e Hans Lukas Teuber. Josiah Macy Jr. Foundation, New York, 1953, 184p.

MACY 10. **Cybernetics***:* circular causal and feedback mechanisms in biological and social systems. Transactions of the tenth conference, April 22, 23 and 24, 1953, Princeton N. J. Editado por Heinz von Förster. Editores Assistentes: Margareth Mead e Hans Lukas Teuber. Josiah Macy Jr. Foundation, New York, 1955, 100p.

MASSARO, L. **Cibernética**: ciência e técnica. 2010. 213p. Dissertação (Mestrado em Sociologia) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas, Campinas, 2010.

McCARTHY, J; HAYES, P. J. Some philosophical problems from the standpoint of artificial intelligence. In: MELTZER, B.; MICHIE, D. (Ed.). **Machine intelligence 4**. Edinburg: Edinburg University Press, 1969. p. 463-502.

McCULLOCH, W. **Embodiments of mind** (collected papers). Cambridge: MIT Press, 1965.

McCULLOCH, W. Mysterium Iniquitatis of Sinful Man Aspiring into the Place of God. **American Association for the Advancement of Science**, v. 80, n 1. p. 35-39, 1955.

McCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115-133, 1943.

MENDONÇA, A. L. O. O legado de Thomas Kuhn após cinquenta anos. **Scientiae Studia**, v. 10, n. 3, p. 535-560, 2012.

MENDONÇA, A. L. O; VIDEIRA, A. A. P. Progresso científico e incomensurabilidade em Thomas Kuhn. **Scientiae Studia**, v. 5, n. 2, p. 169-183, 2007.

MESQUITA, A. P. **Introdução Geral**. Lisboa: Imprensa Nacional-Casa da Moeda, 2005.

MILLER, G.A. **A Very Personal History**. Conferência para o Cognitive Science Workshop, Massachusetts Institute of Technology, Cambridge , Mass., 1º jun. 1979.

MINSKY, M. L **Computation**: finite an infinite machines. New Jersey: Prentice–Hall Inc. Englewood Cliffs, 1967.

MINSKY, M. L; PAPERT, S. **Perceptrons**: an introduction to computational geometry. Cambridge: MIT Press, 1969.

MITCHELL, M. **Complexity**: a guided tour. Oxford University Press, 2009.

NEWELL, A. Intellectual Issues in the History of Artificial Intelligence. In: MACHLUP, F.; MANSFIELD, U. (Ed.). **The Study of Information**: Interdisciplinary Messages. New York, John Wiley, 1983.

NEWELL, A; SHAW, S; SIMON, H. Elements of a theory of GPS. **Psychological Review**, v. 65, 1958.

NEWELL, A; SIMON, H. **Human problem solving**. Prentice–Hall, 1972.

O'REILY, R.C.;MUNAKATA, Y. **Computational explorations in cognitive neuroscience**: Understanding the mind by simulating the brain. MIT Press, 2000.

PENROSE, R. **A mente nova do rei**: computadores, mentes e as leis de física. Trad: Dutra, W. 2ª ed. Rio de Janeiro, 1993.

PESSOA JUNIOR, O. F. O limite Qualitativo de Modelos Quantitativos**. Complexitas – Revista de Filosofia Temática**, v. 1, p. 101-117, 2016.

POPPER, K. R. **A Lógica da Pesquisa Científica**. São Paulo: Cultrix, 1972.

PUTNAM, H. Minds and Machines. In: HOOK, S. (Ed.). **Dimensions of Mind**. New York: New York University Press, 1960.

PUTNAM, H. The meaning of the "meaning". In: PESSIN, A.; GOLDBERG, S. (Org.). **Twin earth chronicles**. New York: M. E. Sharpe, (1996), 1975.

PUTNAM, H. The mental lives of some machines. In: CANTANEDA, H.(Org.) **Intentionality, Minds and Perception**. Detroit, MI: Wayne State University Press, 1967.

ROSENBLATT, F. **Principles of neurodinamics**. New York: Spartan Books, 1962.

RUMELHART, D. E.; McCLELLAND, J. L. **Parallel Distributed Processing**: explorations in the microstructure of cognition. Cambridge: MIT Press, 1986.

RUYER, R. **La cibernetica y el origen de la información**. Ciudad de México: Fondo de Cultura Económica, 1992.

RYLE, G. **The concept of Mind**. Mitchan: Penquin Books, 1949.

SALLES, F. R. **A relevância da cibernética**: aspectos da constribuição filosófica de Norbert Wiener. 2007. Dissertação (Mestrado) – FFLCH-USP, São Paulo, 2007.

SCHANK, R; BIRNBAUM, L. Aumentando a inteligência. In: KHALFA, J. (Org.*)* **A natureza da inteligência**. Trad: Rouanet, L. São Paulo: Unesp, 1997.

SEARLE, J. **Mente, cérebro e ciência**. Tradução de Artur Morão. Lisboa: Edições 70, 1984.

SEARLE, J. **O mistério da consciência**. Tradução de André Uema e Vladimir Safatle. Rio de Janeiro: Paz e Terra, 1998.

SEARLE, J. R. Minds, brains and programs. **Behav. Brain Sci**. v. 3, p. 417-424, 1980.

SHANNON, C.; WEAVER, W. **A mathematical theory of communication**. Urbana: University of Illinois Press, 1998 [1949].

SILVEIRA, F. L. A metodologia dos programas de pesquisa: a epistemologia de Imre Lakatos. **Caderno Catarinense de Ensino de Física**, Florianópolis, v.13, n.3: 219-230, dez. 1996.

SIMON, H. A. **Sciences of the Artificial**. Cambridge, MA: MIT Press, 1969.

SMOLENSKY, P. Connectionist AI, symbolic AI and the brain. **Artificial Intelligence Review**. v.1, p. 95–109, 1987.

SMOLENSKY, P. On the Proper Treatment of Connectionism. **Behavioral and Brain Science**, 11, p. 1-74, 1998.

THAGARD, P. **Mente***:* introdução à ciência cognitiva. Tradução de Maria Rita Hofmesister. Porto Alegre: ArtMed, 1998.

TURING, A. M. Computing machinery and intelligence. **Mind**, v. 59, p. 433-460, 1950.

TURING, A. M. On the computable numbers with an application to the Entscheidungs. In: DAVIS, M. **The undecidable**. New York: Raven Press, 1936.

VARELA, F. J. **Conhecer**: as ciências Cognitivas Tendências e Perspectivas. Tradução de Maria Teresa Guerreiro. Lisboa: Instituto Piaget, 1991.

VARELA, F. J.; THOMPSON, E.; ROSCH, E. **A mente corpórea**: ciência cognitiva e experiência humana. Tradução de Joaquim Nogueira Gil e Jorge de Sousa. Cambridge, MA: Massachusetts Institute of Technology, 1991.

WEAVER, W. Science and Complexity. **American Scientist**, 36: 536. New York City: Rockefeller Foundation, 1948.

WIENER, N. **Cibernética**: ou o controle e comunicação no animal e na máquina. Tradução de Gita K. Guinsburg. São Paulo: Polígono: Universidade de São Paulo, 1961.

WIENER, N. **O conceito de informação na ciência contemporânea**. Rio de Janeiro: Paz e Terra, 1970. (Ciência e informação, v. 2)

WIENER, N.; ROSEMBLUETH, A; BIGELOW, J. Behavior, Purpose and Teleology. **Philosophy of Science** v. 10, n. 1, p. 18-24. Reproduzido (cópia fac-símile) em Masani (1985). Tradução publicada em **Cadernos de História e Filosofia da ciência**, v. 2, p. 43-50, 1943.

Sobre os autores

Alan Rafael Valente

Possui graduação em Filosofia pela Universidade Universidade Estadual do Norte do Paraná - UENP/Jacarezinho (2012 - 2015). Mestre em Filosofia pela Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP/Marília (2016 - 2019). Tem experiência na área de Filosofia, com ênfase em Epistemologia e Filosofia da Ciência, atuando principalmente nos seguintes temas: teoria da informação e da ação, falsificacionismo popperiano e critérios de cientificidade.

João de Fernandes Teixeira

Possui graduação em Filosofia pela Universidade de São Paulo (1977), mestrado em lógica e filosofia da ciência pela Universidade Estadual de Campinas (1982) e doutorado (PhD) em filosofia pela University of Essex (Inglaterra). Fez pós-doutorado nos Estados Unidos em 1998, no Center for Cognitive Studies, na Tufts University, sob a supervisão do Prof. Daniel Dennett. Participou do grupo de ciência cognitiva do Instituto de Estudos Avançados da USP. Foi professor na UNESP (campus-Marília) e da UFSCAR. Publicou 19 livros, a maioria na área de filosofia da mente.

Marcos Antonio Alves

Possui graduação em Filosofia pela Universidade Católica de Pelotas - UCPel (1995), mestrado em Filosofia pela Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP (1999) e doutorado em Filosofia no Instituto de Filosofia e Ciências Humanas/Centro de Lógica, Epistemologia e História da Ciência - CLE - da Universidade Estadual de Campinas - UNICAMP (2012). Professor Assistente Doutor na UNESP no departamento de filosofia e programa de pós graduação em filosofia. Ganhador do 1º Prêmio Marcelo Dascal em Ciência Cognitiva e Filosofia da Mente, tem experiência na área de Filosofia, com ênfase em Epistemologia, Filosofia da Mente e da Informação. Presidente da Sociedade Brasileira de Ciência Cognitiva - SBCC - na gestão 2013-2015, é pesquisador vinculado ao CNPq, líder do Grupo de estudos em filosofia da informação, da mente e epistemologia - GEFIME - e Full member da International Society for Research on Emotions - ISRE. Editor da Trans/form/ação: revista de filosofia da Unesp no biênio 2020-2021 e atual Coordenador do Programa de Pós-Graduação em Filosofia da Unesp.

“Trata-se de trabalho original no qual são articuladas reflexões sobre filosofia da ciência e sobre os fundamentos da ciência cognitiva. O livro aborda tema atual, sobretudo pelo fato de traçar as origens de movimentos importantes na ciência cognitiva, como o conexionismo, que tem servido de base para o *Big Data*.”

João de Fernandes Teixeira

O objetivo desta obra é averiguar o estatuto da ciência cognitiva em sua formação, mais especificamente, entre os anos de 1940 e 1970. Teria ela começado como pré-ciência ou já teria se estabelecido como ciência normal? Para responder a essa questão, os autores apresentam o surgimento dessa área de pesquisa em dois contextos: histórico e epistemológico. Feito isso, analisam, com base no arcabouço conceitual da Estrutura das Revoluções Científicas de Thomas Kuhn, o estatuto científico da ciência cognitiva em sua fase inicial.

ISBN 978-65-5954-051-8
9 786559 540518